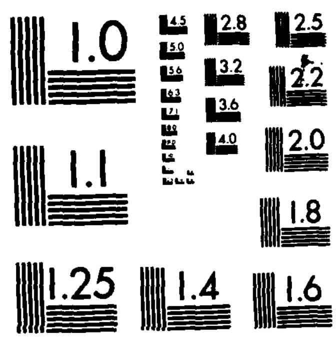
AD-A159 466 MOMENT BY MOMENT VARIATION IN THE COGNITIVE ACTIVITY OF EXPERTS(U) COLORADO UNIV AT BOULDER CENTER FOR RESEARCH ON JUDGMENT AND POLICY R M HAMM AUG 85 CRJP-257 1/1

UNCLASSIFIED N00014-81-C-0591 F/G 5/10 NL

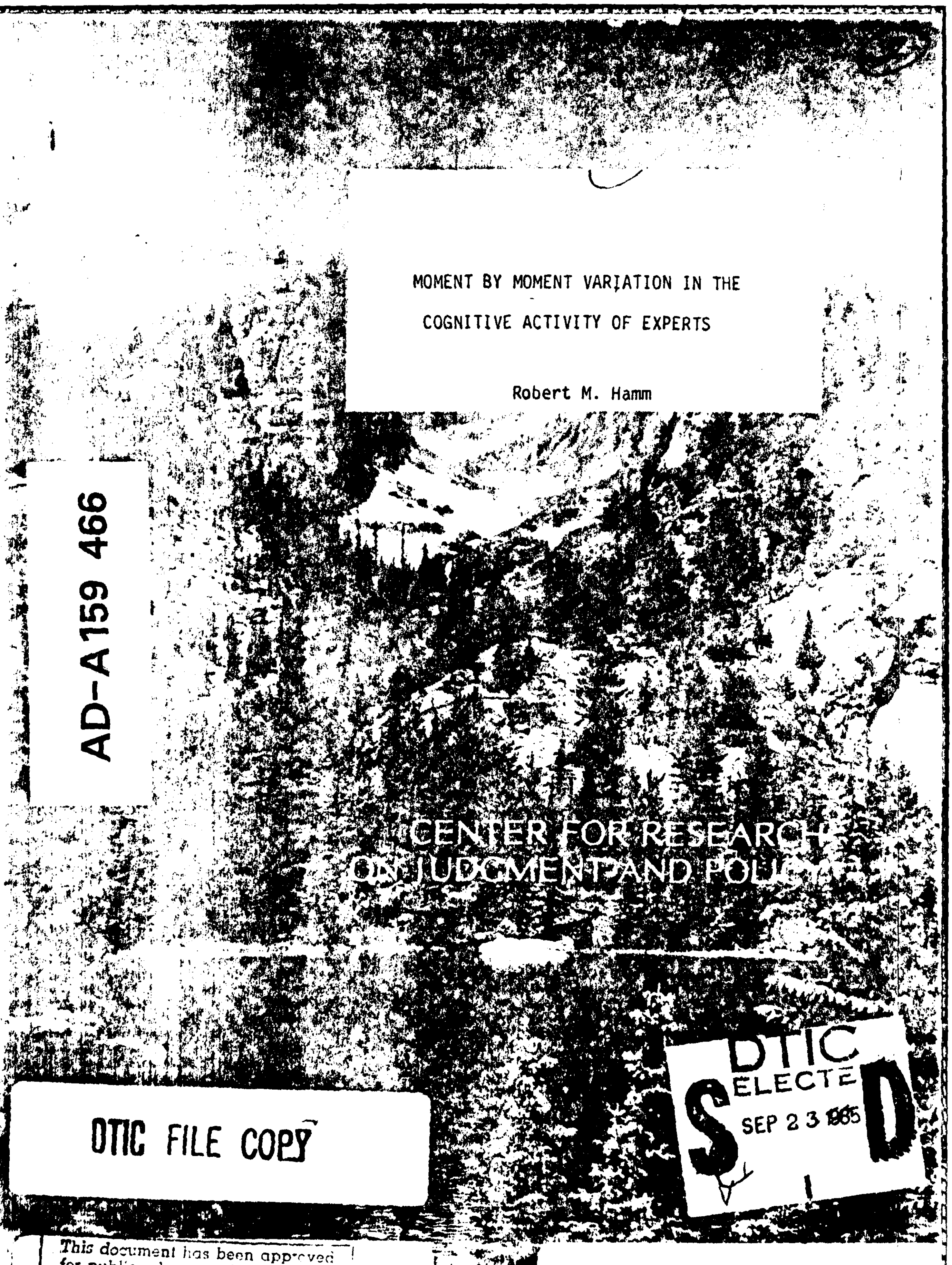
END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MOMENT BY MOMENT VARIATION IN THE COGNITIVE ACTIVITY OF EXPERTS

Robert M. Hamm



MOMENT BY MOMENT VARIATION IN THE COGNITIVE ACTIVITY OF EXPERTS

Robert M. Hamm

Center for Research on Judgment and Policy

University of Colorado

Boulder, Colorado 80309

Report No. 257

August 1985

This research was supported by the Engineering Psychology Programs, Office of Naval Research, Contract N00014-81-C-0591, Work Unit Number NR 197-073 and by BRSG Grant #RR07013-14 awarded by the Biomedical Research Support Program, Division of Research Resources, NIH. Center for Research on Judgment and Policy, Institute of Cognitive Science, University of Colorado.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: CRJP 257	2. GOVT ACCESSION NO.: AD-A159 466
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Moment by Moment Variation in the Cognitive Activity of Experts	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Robert M. Hamm	8. CONTRACT OR GRANT NUMBER(s): N00014-81-C-0591
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Research on Judgment and Policy, Institute of Cognitive Science, University of Colorado, Boulder, CO 80309	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR197-073
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, 800 Quincy Street, Arlington, VA 22217	12. REPORT DATE: August 1985
	13. NUMBER OF PAGES: 82
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public releast; Distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

experts, intuitive-analytical, verbal protocol analysis, Cognitive Continuum Theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper reports the analysis of moment by moment variation in the use of intuitive and analytical cognition by experts engaged in a complex judgment task. Six highway engineers were asked to think aloud while producing formulas that expressed their knowledge of how highway aesthetics, safety, and capacity are determined by sets of relevant factors. Each engineer's statements were measured with multiple indices of the use of intuitive or analytical cognition, thus producing a moment by moment measure of the

20. ABSTRACT (continued)

engineer's cognitive activity. Measures reflecting the rate of alternation between intuition and analysis, as well as linear and nonlinear trends, were derived from this. Differences in these measures due to task topic were found. Two predictions from Cognitive Continuum Theory were supported by the data: (a) the average segment was more analytical on the capacity task than on the aesthetics task, and (b) the rate of alternation between analysis and intuition was faster on the capacity task, which has high standards, than on the aesthetics task. The subtask the engineer was engaged in was also measured over time, and it was found that subjects used different modes of cognitive activity on different subtasks. Finally, it was discovered that the proportion of time the engineer spent doing particular cognitive activities and particular subtasks was related to the production of more accurate formulas.

Moment by Moment Variation in the Cognitive Activity of Experts

The interplay between analysis and intuition has not been extensively investigated by students of judgment and decision making or of problem solving. Psychologists (e.g., Westcott, 1967; Stein, 1974; Hammond, 1980; Isenberg, 1984) and philosophers of science (e.g., Pepper, 1948; Polanyi, 1958) have argued that people's thinking involves both analysis and intuition, and that the two modes of cognition complement each other. As Polanyi put it in discussing mathematical creativity, "The manner in which the mathematician works his way towards discovery, by shifting his confidence from intuition to computation and back again from computation to intuition, while never releasing his hold on either of the two, represents in miniature the whole range of operations by which articulation disciplines and expands the reasoning powers of man" (Polanyi, 1958, p 131). However, there have been no studies of the interplay between intuition and analysis over time. Hence it is not known whether such alternation actually occurs, what factors influence it, nor whether alternation contributes to the quality of thinking.

Hammond (1981) suggests that researchers of judgment, decision making, and problem solving have not dealt explicitly with the complementarity between intuition and analysis because the concepts have not "been securely anchored in theory or in research techniques" (Hammond, 1981, p 3). Hammond (1980, 1981) attempted to provide the theoretical anchor. Work to establish the needed anchors in research techniques has been done by Hammond and colleagues (Hammond, Hamm, Grassia, and Pearson, 1985; Hammond,

Hamm, and Grassia, 1985) and by Howell and colleagues (Howell, 1984; Goldsberry, 1983; Schwartz and Howell, in press). This work took a molar focus. That is, it considered whether a whole task would induce intuition or analysis, and whether the subject's cognition as a whole was intuitive or analytical.

The present research takes a more molecular focus, using a research technique that measures variation in cognition and task on a moment by moment basis. This technique allows the study of changes in cognition over time: Does a person's cognition alternate between intuition and analysis during the execution of a complex task, as Polanyi suggested? What determines the rate of alternation? Further, the measurement of variation in task over time allows the study of the relation between task change and cognitive change. Finally, it is possible to inquire whether the quality of the cognitive performance depends on characteristics of the alternation between intuition and analysis or on the relation between task variation and cognitive variation.

Cognitive Continuum Theory

The research methods of the present paper were invented in order to operationalize and test the account of the temporal alternation between intuitive and analytical cognition provided by Cognitive Continuum Theory (Hammond, 1980). According to this theory, the mode of cognition used on any type of judgment, decision making, or problem solving task can vary on a continuum ranging from intuition to analysis. The cognition in between the two poles, called quasirational cognition, includes elements of both

analysis and intuition. The task partially determines the cognition the person will use on it. Thus, tasks can vary on a continuum, ranging from intuition-inducing to analysis-inducing. The person also has partial control over his or her cognition, accounting for the observation that "cognitive activities move along the intuitive-analytical continuum over time" (Hammond, 1980, p 72). That is, the person can change from intuitive to analytical cognition, or vice versa, even if the given task definition is constant. To account for such changes in cognition, Hammond suggests that "successful cognition inhibits movement, failure stimulates it" (Hammond, 1980, p 72). For example, unfruitful analysis makes likely a shift to intuition.

Testing this theory requires procedures for operationalizing the cognitive continuum and the task continuum. Hammond (1980) listed features of intuitive and analytical cognition. Generally, intuitive cognition is rapid, has low cognitive control (low consistency) and unconscious data processing, uses linear weighted average organizing principles to combine information, etc., while analytical cognition is slow, conscious, and consistent, does not necessarily use linear weighted averaging processes, etc. As for performance, intuitive cognition is predicted to be moderately accurate and moderately inconsistent, while analysis, which is normally highly accurate and consistent, infrequently produces very large errors. The list of features of tasks that induce intuitive cognition includes task unfamiliarity, lack of time, lack of feedback about performance, intercorrelated judgment cues presented simultaneously and perceptually, etc. The opposite features induce analysis. These lists of features

provide the basis for research techniques exploring intuitive and analytical cognition and their dependency on task characteristics.

Previous research on Cognitive Continuum Theory. Hammond, Hamm, Grassia and Pearson (1985) investigated the relation between task and cognition by manipulating both the depth characteristics (e.g., content) and the surface characteristics (information display and response mode) of the task independently, and measuring the cognitive mode the subjects used on the tasks, as well as the accuracy of the subjects' performance on each task. This kind of manipulation varies a number of task features simultaneously. The extent to which each task condition is likely to induce analysis or intuition was measured with a "task continuum index", made by combining measures of eight task features from Hammond's (1980) list. The task features were weighted equally, in the absence of expectations about their relative importance. Similarly, the subjects' cognition was measured on a "cognitive continuum index" which equally weighted measures of four features of cognition from Hammond's (1980) list. It was found that the task partially determines the subjects' cognition, and that their performance is better when their cognition responds to the task's induction. However, the question of cognition shifting over time was not addressed.

Howell and his colleagues (summarized in Howell, 1984) investigated whether task features from Hammond's (1980) lists induce intuition or analysis. In their studies using computer-controlled decision making situations, Howell's group manipulated a total of 14 task features separately (e.g., information load, time constraints) in order to determine

their effects on individual features of cognition (e.g., consistency) and on accuracy of performance. They found that Cognitive Continuum Theory's statements about the effects of task features on cognition are usually but not universally true; that is, the predicted effects depend on other task features or even on the identity of the task. For example, time pressure induced intuition, as Hammond (1980) predicted, only on optional stopping and information integration tasks, but not on the more complicated emergency resource allocation and two-stage judgment tasks. Because the expected task-cognition relations were not found when the task consisted of a number of coordinated subtasks, Howell (1984) suggested that a more "molecular" approach to task description is needed.

Foundations of the Present Research: New Measures

The study to be reported below used such molecular measures to address issues concerning the relation between task and cognitive mode over time, as well the issue of temporal changes in cognitive mode.

Molecular measures of task. A more molecular task description technique requires measurement of task features at a particular juncture in the subject's judgment or problem solving process, rather than measurement of the features of the task as a whole. Such time-specific task descriptions would allow for more precise tests of Cognitive Continuum Theory's predictions about the relation between task features and cognitive mode. But more molecular measures of *cognitive mode* would be needed, as well.

Molecular measures of cognition. The methods previously used for measuring cognitive mode have taken a "molar" time perspective. They have addressed the whole of the subjects' performance, whether the focus has been on one feature (Howell, 1984) or on the average of several features (Hammond, Hamm, Grassia, and Pearson, 1985). For example, the consistency parameter from the Brunswik Lens Model (used by both Howell and Hammond) and the peakedness of the distribution of a subject's judgment errors (used by Hammond) are summary concepts derived from data about a large number of the subject's judgments. These measures are not appropriate for tracing momentary changes in the subject's cognitive mode. However, other features from Hammond's (1980) list may be ascertained on the required moment-by-moment basis, and so were used in the present study.

Foundations of the Present Research: Theory Concerning the Relationship over Time between Task and Cognition

Despite the absence of research techniques sensitive to change over time, temporal changes in cognitive mode play a prominent role in Cognitive Continuum Theory. The fourth of Hammond's (1980) five premises holds that the cognitive mode changes over time, presumably under the control of the subject rather than of changes in the externally defined task. Cognition is expected to change mode when the current mode is not successful, a sequential or temporal phenomenon. Two types of quasirational cognition were also distinguished in terms of their temporal characteristics: when the task demands high performance, cognition will shift between analysis and intuition, but when the task will accept "good enough" performance then

there will be cognitive "compromise", i.e., the steady use of an intermediate mode of cognition. Payne (1982) suggested that cognition can undergo development from intuition to analysis. Such a process could occur on either the macro level (e.g., maturation, the attainment of expertise) or the micro level (starting a task intuitively and moving to analysis). Each of these ideas involves only changes of cognition, and assumes a constant task. Such changing cognition could account for Howell's (1984) finding that time pressure did not influence cognitive mode in the predicted manner on the complicated tasks. If the subjects' cognition undergoes development from intuition to analysis on these tasks, or if the task demands high performance and the subjects consequently shift cognitive mode rapidly, such changes in cognition would be lost to observation when the only available measure is derived from the features of cognition on the overall task.

We must recognize, further, that it is often the subject rather than the situation (the researcher) who defines the task he or she is doing at any given moment (Payne, 1982). For example, Kahneman and Tversky (1979) described how subjects choose among two gambles by adopting a two-phase strategy. The first phase involves editing or framing the problem, a process Payne suggested is intuitive, for it involves perception-like processes such as pattern recognition. The second phase, evaluating the gambles within the chosen framework, would be analytical (Payne, 1982, p 399). By this view, changes in cognitive mode would be due to changes in the subject's self-defined task. Evaluating a risky choice would involve moving from intuitive cognition to analytical cognition: first one sets

oneself an intuition-inducing subtask (framing) and then one sets oneself an analysis-inducing subtask (evaluating the prospects within the framework). (See also Hammond's (1981) discussion of the Maier Two-String Problem.) Any arbitrary pattern of shifts in cognition might occur on a decision problem, depending on the order of the subtasks the subject engages in and on how analysis-inducing each subtask is.

Finally, even though the subject may have defined the task, the task does not necessarily completely determine the subject's cognition. For example, the subject's cognitive mode might undergo development on the micro level within each of the subtasks he or she sets for self. Upon successful completion of a subtask, the subject would start a new subtask intuitively, and slowly get more analytical until the subtask was complete. The next subtask would be started intuitively again. This account would entail shifts from analysis to intuition upon success, the opposite pattern from Hammond's (1980) prediction of shift upon failure.

In order to study these conjectures about the variation in cognitive activity and its relation to the task, research techniques measuring the extent to which cognition is analytical or intuitive on a molecular time basis are needed, as well as techniques for analyzing the subject's subtask as it changes moment by moment. To develop the required techniques, a study was carried out in which task features and cognitive mode features were measured on a moment by moment basis while subjects engaged in a complicated judgment task.

The Judgment Task Used in the Present Study

Producing a mathematical formula for expressing one's knowledge of the relations among entities is a commonly exercised judgment skill. For example, teachers may invent formulas for determining students' course grades from their scores on a number of tests, assignments, and papers. People use similar formulas for rating the quality of their available options when making decisions, using the techniques of decision analysis (e.g., Keeney and Raiffa, 1976; Edwards and Newman, 1982; Behn and Vaupel, 1982).

Besides using formulas to express judgments of quality or preferability, such procedures can be used in externalizing experts' implicit knowledge. For example, the procedure known as "adaptive environmental assessment" convenes ecology experts for workshops focussed on particular topics, during which the experts' knowledge of particular relations is expressed in formula form, and then these formulas are combined into an overall model of the ecological problem (Holling, 1978). Hammond, Anderson, Sutherland, and Marvin (1984) used a related procedure, in which the experts were consulted individually instead of in group format, to assess the health risk of plutonium emissions from a weapons factory. Kirwan, Chaput de Saintonge, Joyce, Holmes, and Currey (1985) required rheumatologists to produce formulas for predicting the severity of patients' illness as a standard for comparison with other forms of self-report. Finally, though the possibility does not seem to have been recognized, such techniques would also be applicable to the problem of extracting intuitive knowledge from experts to build expert systems (e.g., Buchanan and Shortliffe, 1984; Davis, 1979).

Because of the range of interest in experts externalizing their knowledge of a domain using mathematical formulas, Hammond, Hamm, Grassia, and Pearson (1985) studied individual highway engineers' use of mathematical formulas to express their knowledge of the dependence of one feature of highways (such as highway safety) upon sets of relevant factors (such as the width of the lanes and the shoulders, the frequency of curves, intersections, and roadside obstacles, the volume of traffic and the proportion of trucks.) Verbal transcripts collected during the production of these formulas makes it possible to study variation in cognitive mode during such activity. Six engineers were asked to think aloud while producing three formulas each. The present study analyzed their thinking aloud protocols, measuring the presence of features of intuitive and analytical cognition from Hammond's (1980) list at each moment. This makes possible the calculation of a moment by moment index of the engineer's position on the cognitive continuum (MBMCCI). In addition, although the researchers defined only a single task for the engineer--to produce a specified formula--nonetheless, the subject may be expected to engage in a number of different self-defined subtasks in the course of producing a formula. The subtask the subject was engaged in was registered at each moment.

Problems Studied

These "molecular" or moment by moment measures of cognitive mode and task make it possible to address the following issues in this study:

Problem 1. Stability and shifting of cognitive mode. Formula production is largely an analysis-inducing task, yet the "alternation hypothesis" predicts that the subject's cognition will vary between intuition and analysis. The MBMCCI should reveal evidence of shifts over time in the engineer's mode of cognition. It is not known whether these changes will be rapid alternations (Curve 1), steady linear trends (Curves 2 and 3 in Figure 1), nonlinear trends (Curves 4 and 5), or some combination of these patterns. Procedures that will be developed for measuring the degree to which each of these patterns occurs will allow comparison of the patterns.

Insert Figure 1 about here

Problem 2. The relation of the pattern of change in cognitive activity to the engineer's performance. The accuracy of these subjects' formulas was evaluated in the previous study (Hammond, Hamm, Grassia, and Pearson, 1985). This information can be used to explore whether the patterns of variation in the MBMCCI (Figure 1) are related to higher accuracy in the formulas.

Problem 3. The influence of the stringency of task standards on the pace of shifting among cognitive modes. Hammond (1981, p 44) predicted that subjects' quasirational cognition will shift between intuition and analysis when the task has high standards, while staying more steadily at an intermediate level when the task has low standards. Although the three formula—making tasks were presented identically in this study, the engineers had lower expectations for their aesthetics formulas. Engineers

are traditionally responsible for judgments of highway safety and capacity, but not for aesthetics judgments. Further, there is a known procedure for calculating capacity which many of our engineers referred to explicitly when making their own formulas. Finally, a number of the engineers expressed doubt or amusement at the task of making an aesthetics formula. Since they expected less from their aesthetics formulas, their rate of shifting between analysis and intuition should have been less than on the capacity and safety tasks.

Problem 4. The influence of the topic of the formula task on cognitive activity. Each subject wrote formulas for three different highway concepts: aesthetic value, safety, and vehicle-carrying capacity. It was demonstrated by Hammond, Hamm, Grassia, and Pearson (1985) that the engineer's cognitive mode varied as a function of the topic of the formula--aesthetic value was the most intuition-inducing concept, while capacity was the most analysis-inducing. However, in that study cognitive mode was measured in terms of features of the overall judgment policy (i.e., its consistency, its linearity, the subject's confidence in it, and the distribution of its errors). The MBMCCI, averaged across all the segments in a formula making session, is a conceptually related but operationally distinct measure that provides a separate test of Cognitive Continuum Theory's prediction that the aesthetics formula task will induce more intuitive cognitive activity than the capacity task, with the safety task in between.

Problem 5. The relation between subtask and cognitive mode. Each of the measures of moment by moment changes in the engineers' self-defined subtask allows testing of whether the engineers' cognitive mode shifts according to the kind of subtask.

Method and Procedure

The subjects were six male highway engineers (ages 30 to 54), a subset of the 21 engineers used in Hammond, Hamm, Grassia, and Pearson (1985). The engineers were asked, in separate sessions each lasting an hour or more, to construct mathematical formulas expressing their knowledge of how the safety, aesthetic value, and vehicle-carrying capacity of 2-lane rural Colorado highways depends on sets of relevant variables. The order in which the engineers produced the three formulas was counterbalanced. (The transcript of Subject #1's aesthetics formula session was lost and so only 17 sessions were coded and analyzed.) The formula-making tasks were the last three of nine tasks done by each engineer in this study. Thus, he had seen filmstrips of 40 highways representative of the class for which he was writing a formula, and the researcher had twice before guided him in thinking about aesthetic value, safety, and capacity, once using the sets of relevant dimensions that he would be asked, in these sessions, to use in his formulas (Hammond, Hamm, Grassia, and Pearson, 1985).

The engineers produced the formulas in their offices. The engineer was asked to schedule an hour and a half for each session. His task was defined by an instruction sheet (Appendix A) which asked him to "develop a general procedure for making a judgment of the aesthetic value [or safety

or capacity] of highways. That is, you will make up a formal procedure, such as a mathematical equation, that will take a highway, described in terms of its important characteristics, and produce a number that reflects its aesthetic value [or safety or capacity] as accurately as possible." The engineer was allowed to use pencil and paper, graph paper, ruler, and calculator.

The think-aloud instructions were on a separate sheet (Appendix B) which asked the engineer to tell the researcher "EVERYTHING you are thinking from the time you are first given the next sheet of instructions, until you are satisfied with the formula you have produced." This procedure, a variant on instructions used by Anders Ericsson (personal communication), included a practice exercise. The tape recording started with the engineer reading the definitions of the dimension the formula was meant to predict, e.g., aesthetic value (Appendix C), and of the dimensions to be used in the formula (Appendix D).

Analysis of the accuracy of the formulas and of the kind of formula produced (in terms of whether the organizing principle was additive or multiplicative, whether tables were used, etc.) revealed no systematic differences between the six engineers who produced their formulas while thinking aloud and the other 15 engineers in the study.

To facilitate coding of both task and cognition on a molecular or moment by moment basis, each engineer's thinking aloud was recorded, transcribed, and segmented into units that were numbered sequentially from when he began reading the dimension definition (Appendix C) until he said

his formula was finished. Segmentation into units was guided by the principle that there be only one kind of cognitive activity in each segment. Thus, the transcript was broken into short sentences, clauses within longer sentences, and incomplete sentences (see Figure 2). Each segment was coded with respect to nineteen different categorization schemes relating to the engineer's cognitive mode and self-defined task. The categorization schemes were divided among four different coders (see Table 1). Reliability was measured by having a second coder duplicate the coding on three transcripts.

Insert Figure 2 and Table 1 about here

Fourteen of the coding schemes were designed to measure activities that are indicative of analytical or intuitive cognition. Four measured the self-defined task the engineer was engaged in. The final categorization scheme identified whether the engineer was thinking concurrently or retrospectively, a distinction drawn by Ericsson and Simon (1984). Segments were rarely coded as "retrospective". These occurred mainly after the researcher prompted an engineer who had fallen silent, and were dropped from subsequent analyses.

Coding judgments were not made strictly independently. First, coders were instructed to use contextual information to help them interpret each segment. While using context increases coding accuracy, the coding of one segment is not independent of the coding of its neighbors. Second, each coder applied several categorization schemes in parallel. There may be

halo effects from one categorization scheme to another, within coder. The categorization schemes coded by each coder are given in Table 1. Two of the coders (RH and JG) were not blind to the hypotheses or to the fact that the codes were to be used as subindices of a cognitive continuum index. These coders coded only features of cognition, and thus their knowledge could not influence any relations subsequently shown between the indices of cognition and the identification of self-defined tasks.

Coding for Features of Analytical and Intuitive Cognition

The strategy of the MBMCCI method is to measure those features reflecting intuitive or analytical cognition (Hammond, 1980) that can be identified in individual segments of the transcript. However, the features are not perfect indicators of the thinker's position on the cognitive continuum; for example, in one segment there may be features characteristic of intuition and of analysis, or the relation of a feature to analysis or intuition may be ambiguous. Because of this imprecise measurement, it was decided to measure fourteen features or activities reflecting analytical or intuitive cognition, and to average them to produce an overall index of the engineer's position on the cognitive continuum during a particular segment.

The categorization schemes are: (A) Decisions; (B) Justifications; (C) Memory use; (D) Source of knowledge; (E) Kind of knowledge of correlation; (G) Kind of judgment; (H) Degree of quantification of a judgment; (J) Difficulty verbalizing; (K) Confidence and doubt.[1] The categories in each scheme are given in Table 2. For categorization schemes A to E, G, and H, each segment could be given only one category from each

scheme, so the coder had to choose the most appropriate one if several were applicable. If no categories were applicable, the segment was coded as "irrelevant" with respect to the scheme. For schemes J and K, the number of instances of each category in each segment (e.g., "pauses" in scheme J) was counted; a score of 0 was not considered "irrelevant".

Insert Table 2 about here

The Moment by Moment Cognitive Continuum Index

To produce the MBMCCI index, scales were constructed for each categorization scheme and these scales were combined in a weighted average procedure. To produce the scales, the categories in Schemes A to E, G, and H were ranked according to the author's interpretation of the level of analysis that each of the categories would entail. For example, taking an action without justification was considered the most intuitive decision making category, and considering tradeoffs among dimensions of outcomes was considered the most analytical. (See the Coder's Manual (Hamm, 1985) for more detailed definitions of the categories.) Integers were assigned to each category (see Table 2), starting with 0 as the most intuitive. Finally each scheme's scores were divided by the score assigned to the most analytical category in the scheme, producing a scale that ranges from 0 (most intuitive) to 1 (most analytical). For example, a segment's scale value on Scheme A would be A/6, where A is the category score (from Table 2) and 6 is the score assigned to the highest category in the scheme. Schemes J and K were scored differently. The original coding was the count

of how many instances of the category occurred in the segment. This number was used directly as the score for the category.

The scales from each of the schemes were averaged, using the following weights:

Schemes	Weight
A,D,H	2
B,E,G,Jre,Jch,Jinc,Kdbt,Kcnf	1
C,Jpa,Jmut	0.5

The source of these weights was the author's a priori judgment of the schemes' relative quality as measures of cognitive mode. In sum, the MBMCCI for each segment was calculated by the following formula:

$$\text{MBMCCI} = \text{mean}\left[2\left(\frac{A}{6}\right),\ B,\ .5\left(\frac{C}{2}\right),\ 2\left(\frac{D}{3}\right),\ \frac{E}{3},\ \frac{G}{2},\ 2\left(\frac{H}{2}\right),\ .5\left(1-\frac{Jpa+Jmut}{2}\right),\ 1-\frac{Jre+Jch+Jinc}{3},\ (1-Kdbt),\ Kcnf\right]$$

where the letters represent the category scores indicated in Table 2, and the "mean" operation excludes schemes that were irrelevant. This measure can range from 0 (if every subindex is maximally intuitive) to 1. It could exceed this range only if there were counts larger than 1 for the segment on the J and K categories. If none of the categories in any scheme were relevant, and if there were no instances of the categories in Schemes J or K, the MBMCCI score would be .417, because: Schemes B, E, and Kcnf count as 0's; Schemes J and Kdbt count as 1's; the other schemes make no

contribution; and using the weights in the formula yields a mean of .417. In this way, the MBMCCI measure avoids defining segments as 'intuitive' simply because of lack of evidence of analysis.

The subindex codes and the MBMCCI for each segment in Figure 2 are given in Table 3. A score of "ir" meant that no category in the scheme was relevant for the segment. Some segments are missing, on the basis of being coded as an unfinished remark by coder RL. One of the most intuitive segments is #277, "As you get 7 to 10 feet with 5 [objects per mile]", which has an MBMCCI score of .32 because the engineer is not revealing any conscious decision making about what he is doing now (Scheme A), is making a statement without justification (Scheme B), is using memory implicitly in making a judgment (Scheme C - the judgment is stated in segment #280), based on his experience (Scheme D), is not referring to any causal or correlational connections among dimensions (Scheme E), and is speaking only in terms of the qualities of the dimensions he is considering (Scheme G), although he <u>does</u> use numerical quantification (Scheme H). Though there is repetition of "5" between segments 277 and 278 (Scheme Jre) and there is no expression of confidence (Scheme Kcnf), the other signs of intuition (difficulty verbalizing and doubts) were lacking.

Segment #291, "probably by 50 percent," was tied for most analytical in this excerpt (MBMCCI = .53) because the engineer was judged to be conscious of the fact that he was making a choice about how to proceed (Scheme A) and was using causal knowledge in a predictive manner (Scheme E) to state a relationship (Scheme G) using numerical ratings (Scheme H).

Insert Table 3 about here

Coding of the Engineer's Self-defined Tasks

Four categorization schemes were used for coding variation in the task the engineer defined for himself at each stage in the production of a formula (see Coder's Manual for more detail). Each scheme addresses a different informal model of formula production.

Decomposition (Scheme F). This informal model captures the notion that a complicated task is approached analytically by breaking it into parts and doing each part in turn (cf. Anderson, Deane, Hammond, McClelland, & Shanteau, 1981, p 79). The categories of this coding scheme are: name/register the goal, break it down into parts, structure the subtasks, perform the subtasks, combine the results of the subtasks (see Table 4).

Formula parts (Scheme I). This informal model assumes that the subject performs subtasks particular to the writing of formulas: thinking about the whole formula, the organizing principle, or the dimensions (cf. Anderson et al.,1981, p 6).

Search (Scheme II). This informal model holds that one solves a problem by searching through a space of possible solutions (Newell and Simon, 1972; Bradshaw, Langley, and Simon, 1983). The categories in this coding scheme are: information gathering, pregenerational activity, constraint setting, generation of formulas or formula parts, evaluation of formulas or formula parts, report of product.

Information Processing (Scheme III). This informal model has its source in the cognitive psychologists' idea that people use many information processes (Newell and Simon, 1972; Lindsey and Norman, 1972), which we may consider to be "subtasks". It identifies each segment as involving control (making a plan, using a plan, evaluating the outcome), memory (storing, recalling, using mental imagery), or judgment (verbal, numerical).

The subtask categories for the segments in Figure 2 are presented in Table 5. These coding schemes are used to identify the self-defined subtasks, not to measure how analysis-inducing or intuition-inducing the subtasks are. Thus each subtask categorization scheme will be used independently in the analysis.

Insert Table 4 and Table 5 about here

Analysis and Results

The results will be presented as follows. First the characteristics of the moment by moment index of the position on the cognitive continuum of the engineer's cognitive activity will be described--its distribution, reliability, and stability. The next topic is Problem 1, the patterns of variation in the MBMCCI index over time; measures of the rate of alternation on the micro level, and linear and nonlinear trends on the macro level, will be described. Problem 2, the relation of the patterns of temporal variation to the accuracy of the formula, is addressed next. Problems 3 and 4 are concerned with testing specific predictions from

Cognitive Continuum Theory. Then the results of the coding of the subtask categorization schemes are described. Problem 5 asks whether there are differences in the engineer's cognitive activity on the different subtasks. The final section explores the relation of the formulas' accuracy to the various activities that serve as measures of the MBMCCI and to the engineers' use of the various subtasks.

Preliminaries: The Moment by Moment Cognitive Continuum Index

The MBMCCI can range from 0 (most intuitive) to 1 (most analytical). The mean MBMCCI was calculated for each session, and ranged from .342 to .437, with a mean over all 17 sessions of .388. The standard deviation of the MBMCCI ranged over sessions from .086 to .132, with a mean of .110. Two very analytical segments from a typical session (Engineer 15, capacity formula) are "one or both sides of the road" (MBMCCI = .89) and "so I can deal in some common units here" (.85). Two very intuitive segments are "the thing that's confusing about all this is" (.18) and "and the average speed limit is not--not really very important here" (.17).

Reliability of the subindices. Each scheme was coded throughout all 17 sessions by a single coder, and on three sessions by a second coder to determine the reliability of the coding (see Table 1). The reliability of each category in Schemes A to E, G, and H is determined by (a) creating a variable on which a segment is scored as "1" if the category were assigned to it and "0" otherwise, and (b) correlating the corresponding variables for the two coders. The median reliability (of the three sessions that were checked) for each category is given in Table 6. To determine the

reliability for categories J and K the two coders' counts for each category were correlated. If a category was not used by one of the coders, no reliability correlation could be computed. In Table 6, a "~" indicates that there was no reliability correlation for any of the three reliability sessions. Reliabilities ranged from -.03 for the 'comparison' category of Scheme A to .97 for the 'pauses' category of Scheme J. More details are given in the Coder's Manual (Hamm, 1985).

Insert Table 6 about here

The reliability of Schemes A to E and G to K with respect to the measurement of cognitive mode was measured by (a) creating a variable which assigned to each coder's category judgments the designated scale values (Table 2), and (b) correlating the corresponding variables for the two coders (Table 6). The median scheme reliabilities (of the three sessions that were checked) ranged from -.02 for Scheme E to .86 for Scheme H. No scheme was dropped from the index due to unreliability.

The reliability of each MBMCCI subindex coder was determined by (a) producing the subset of the MBMCCI that could be calculated from each coder's schemes and (b) correlating the corresponding indices for the two coders. The median reliabilities are given in Table 1: .52 for Schemes C, D, E, G, and H, .42 for Schemes A and B, and .69 for Schemes J and K. Since the MBMCCI is a weighted mean of these three scores, its reliability may be estimated, using the Spearman-Brown formula (which assumes equal weights), as .78.

Stability of cognitive activity over time. Before considering whether the subjects' cognitive activity alternates over time between intuition and analysis, it is necessary to ask whether there is evidence for stability of the mode of cognitive activity, as opposed to a random pattern of variation in the MBMCCI. Two analyses of the MBMCCI's variation over time are pertinent to the stability question: autocorrelations and runs.

Autocorrelations. The correlations between the MBMCCI scores of neighboring segments reflect their temporal stability. The MBMCCI measures of the segments in the transcript constitute a series of observations over time. Autocorrelation is the correlation of a time series variable with a second variable whose values are the value of the first variable one or more time periods earlier. Thus, an autocorrelation at 1 lag is the correlation of the time series variable with itself one period earlier. The autocorrelations at lags 1 through 25 were calculated for each session. The mean autocorrelations at lags 1, 2, 3, and 4 were .40, .27, .17, and .13, respectively. The first lag with a negative autocorrelation ranged from the third to something greater than the 25th, with a median of the ninth. The mean number of positive autocorrelations out of the first 25 lags was 16, and ranged from 8 to 25. These findings indicate that there is continuity between neighboring segments.

Runs. A second indicator of stability is the length of runs of segments with similar cognitive activity. If the segments within each session are dichotomized into an intuitive and an analytical group, the number of times the engineer shifted between analysis and intuition (i.e.,

the number of runs) is a measure of the stability of the engineer's cognition. If there were no segment-to-segment continuity, that is, if the MBMCCI reflected a random process, then for N segments of which K are in the analytical group, the expected length of the runs would be (N**2)/2*K*(N-K) or 2 if there are equal numbers of analytical and intuitive segments.[2] The "intuitive" and "analytical" groups were defined using either the mean or the median of the MBMCCI scores in the session as the cutpoint, whichever produced a partition with more equal numbers of segments in each group. The mean length of run in a session was 2.9 segments, ranging from 2.3 to 5.2; this is longer than expected for every session. The runs test determines the likelihood of such a finding by calculating a Z score representing the probability that a given number of runs would be observed if there were truly no continuity between the neighboring segments. If the Z score is negative, it means that there are fewer (longer) runs than chance expectation, i.e., that there are fewer shifts between analysis and intuition. The Z scores for this test ranged from -1.97 to -11.08, with a mean of -4.58, showing that each session had significantly fewer shifts than would be expected by chance. This is evidence for stability in the engineers' cognitive activity.

Both the autocorrelation and the runs techniques thus found stability in cognitive activity on the moment by moment level. This does not mean that there is _no_ change over time; rather, more _credence_ can now be given to observed changes in MBMCCI, since the variation is non-random. These non-random changes in cognitive activity can be studied on both a micro

level (the rate of alternation on the segment to segment scale between intuition and analysis) and a macro level (trends on the scale of the whole session).

Convergent Validity. The correlation of the mean MBMCCI with the CCI measured on these sessions in the other study (Hammond, Hamm, Grassia, and Pearson, 1985) is .27 (N = 17, NS). However, the four subindices of the CCI measure were scaled independently for each engineer, so they are not really comparable. If the subindices are put on common scales before being averaged, a "comparable CCI" is produced which correlated .67 with the old CCI. The "comparable CCI" correlates only .07 with the mean MBMCCI. Though the two concepts have the same name, one is measured very differently from the other, so lack of correlation is in a sense not surprising (see Discussion).

Problem 1: Patterns of Change in the Engineers' Cognition over Time

The MBMCCI provides a measure of the degree to which the engineer's cognitive activity is intuitive or analytical at each moment during the performance of a complex task. It is possible to study changes in the MBMCCI on two levels. On the local or micro level, the rate at which the engineer shifts between analysis and intuition can be measured, and on the global or macro level, linear or nonlinear trends from one type of cognition to the other can be identified.

Change at the micro level: Rate of alternation between analysis and intuition. The autocorrelation and runs techniques provide several indices which reflect not only the degree of stability but also its inverse, the degree or rate of alternation. A measure of the rate of alternation at the micro level (on the scale of up to 25 segments at a time) can be produced by averaging standardized measures from the autocorrelation analysis (mean autocorrelation of lags 1 to 4; number of positive autocorrelations out of lags 1 to 25) and from the runs analysis (size of the Z score, mean run length). The subindices from the two procedures are highly intercorrelated (see Table 7) and so they have high convergent validity. This index reflects the existence of patterns like Curve 1 in Figure 1, and it can be used for comparing the rate of alternation of such patterns across sessions. For example, Column 2 of Table 8 shows that Engineers 1 and 10 had slower rates of alternation between intuition and analysis than the other 4 engineers (one-way anova, $F(5,11)=2.85$, $p = .068$).

Insert Table 7 and Table 8 about here

Change at the macro level: Linear and nonlinear trends over the whole session. The above measure of micro alternation rate does not exhaust the concept of changes between intuition and analysis over time. The components of the alternation rate measure, particularly those derived from the autocorrelation analysis, refer to moment by moment changes in cognition. There may, however, be changes in cognition on a larger scale, e.g., a steady linear trend toward analysis (or toward intuition), or a

nonlinear trend from intuition to analysis and back to intuition (or vice versa). While the size of the autocorrelations generally decreased with increasing lag (autocorrelations up to lag 25 were inspected), and no significant negative autocorrelations (indicating a <u>regular</u> alternation) were observed, it is still possible that there were patterns of change on a much larger time scale - trends over the course of the whole session. To investigate this possibility, each session was divided into tenths and one-way analyses of variance were performed for each session separately to test for (a) a steadily increasing or decreasing trend and (b) a nonlinear change pattern, i.e., a trend in one direction during the first half of the session which reverses direction during the second half.

<u>Linear trends: Steady increase or decrease in analysis over time</u>. There was no evidence for a general linear trend between the beginning and end of the sessions. Ten of the seventeen sessions showed trends toward increasing analysis, and the remaining seven showed a trend toward increasing intuition, but the effects were small. However, a third of the sessions manifested linear trends to a statistically significant degree. One safety formula session ($p < .01$) and three aesthetics sessions ($p < .05$) became increasing analytical, and two safety sessions (one $p < .05$ and the other marginal at $p < .10$)) became increasingly intuitive.

The direction and intensity of this linear trend can be measured by assigning to its R**2 (the proportion of the session's MBMCCI variance that the trend accounts for) a positive sign if the direction of the trend is toward analysis or a negative sign if the direction is toward intuition. This linear trend index ranged from -.01 (signifying that only 1% of the

MBMCCI variance in the formula-making session is accounted for by the engineer's cognitive activity becoming increasingly intuitive as the session progressed) to .18 (18% of the variance being due to the engineer's increasing use of analysis). The mean trend index was .016, and the mean unsigned R**2 (indicating the magnitude of the linear trend irrespective of its direction) was .019. Therefore there is no general trend of changing cognitive activity over all sessions, and even in the exceptional individual sessions the trend accounted for only a small proportion of the MBMCCI variance. Neither the formula tasks, the ordinal position in which the engineer produced the formula, nor the identity of the particular engineer (Column 3 of Table 8) had any effect on the index of linear trend when evaluated with one-way analyses of variance ($N = 17$; probability levels are given in Table 8). Repeated measures analyses of variance, which are more appropriate for testing the formula task and ordinal position effects, yet which do not allow explicit comparison of the subjects, also showed insignificant task and order effects (see Column 3 of Table 8).

Nonlinear trends: Cycles on the scale of the whole session. There was evidence for nonlinear trends over the course of each session, although the shape of the trend varied. The patterns were identified using contrasts in the one-way analysis of variance of the individual sessions. In thirteen of the sessions the engineers started with relatively analytical cognition, became more intuitive up to the middle of the session, and then became more analytical again (A-I-A). The contrast for the A-I-A pattern was significant for four aesthetics sessions (three at

the $p < .001$ level and one at the $p < .01$ level). The four remaining sessions (three of them capacity sessions) showed the opposite nonlinear trend (I-A-I). The contrast for this pattern was significant for two capacity sessions (one at $p < .001$, the other at $p < .05$) and for a safety session ($p < .05$).

The direction and intensity of this nonlinear trend is measured by assigning to its R**2 a positive sign if the pattern of alternation is I-A-I and a negative sign if it is A-I-A. This ranged from -.40 (signifying that 40% of the variance of the MBMCCI score is due to the A-I-A pattern) to .10 (10% due to the I-A-I pattern). The mean score on this signed index was -.05, and the mean unsigned R**2 was .064. Hence the general pattern was A-I-A, though there were substantial individual session differences.

The aesthetics sessions showed the A-I-A pattern very strongly. All 5 engineers had A-I-A rather than I-A-I, and a mean of 19% of the MBMCCI variance was accounted for by this pattern (Column 4 of Table 8). The differences among task sessions in the strength and direction of the nonlinear trend, due largely to the strong aesthetics A-I-A pattern, was significant at $p = .002$ (one-way analysis of variance).

<u>Comparison of the linear and nonlinear trends</u>. The nonlinear trend accounted for an average of 6.4% of the variance, as compared with the linear trend (only 1.9%). Further, there were differences among the sessions in the direction of the nonlinear trend, for there were sessions that had the opposing I-A-I and A-I-A patterns, each at the $p < .001$ level.

In sum, this analysis of change in the MBMCCI over the whole session shows evidence for a pattern of alternating between intuition and analysis (both Curve 4 and Curve 5 of Figure 1), but no evidence for a pattern of steadily becoming more analytical or more intuitive (Curves 2 and 3 of Figure 1). The trend indices, together with the index of alternation rate (Curve 1 of Figure 1), provide measures for describing the changes between analytical and intuitive cognitive activity over time. Such tools have not previously been available, and allow new classes of questions to be investigated. For example, do these change patterns determine the quality of the engineers' performance?

<u>Problem 2: The Relation of Formula Accuracy to the Measures of Change over Time</u>

The achievement (r_a) measure of the accuracy of the engineer's formula (the correlation between the answers his formula produced, when applied to a set of highways, and a criterion measure) was used to determine whether features of the engineers' alternation between intuition and analysis (the alternation rate and the linear and nonlinear trends) influence accuracy. Each engineer's achievement was calculated by Hammond, Hamm, Grassia and Pearson (1985) and ranged from -.57 to .95 with a mean of .52. The correlations between achievement and the measures of alternation rate and linear and nonlinear trend are -.06, -.17, and -.41, respectively (Table 9), which are not statistically significant ($N = 17$). Further analysis, however, reveals that the relations vary for the different formula tasks (Table 9). Thus, the slope of the relation between achievement and

alternation rate is -.25 (correlation = -.72, $p < .10$) for the safety sessions, indicating that the engineers who switch more slowly between analytical and intuitive cognition write less accurate formulas. But there was no relation for the aesthetics formula (slope = -.01) and the capacity formula showed an inverted U-shaped curve -- sessions in which there was a moderate rate of alternation produced the best formulas (its linear slope is .11).

Figure 3 plots these data and shows the slopes for the linear relation between achievement and alternation rate for each of the formula tasks. In an analysis of variance of the effect on achievement of the formula task, with the alternation rate as a covariate, the alternation rate by formula task interaction tests whether the slope of the linear relation between alternation rate and achievement is significantly different for the different formula tasks. This interaction term was nonsignificant, as was the overall effect of alternation rate. However, further analysis taking into account the nonlinear relation of achievement to the alternation rate of the capacity sessions might reveal this effect to be statistically significant. Similar suggestions of formula task effects on the relations of linear and nonlinear trends to achievement may be seen in Table 9 and are displayed in Figure 4 and Figure 5, respectively. The overall relation between the nonlinear trend and formula accuracy nears significance ($F(1,11)=3.25$, $p = .099$), indicating that the more the engineer's cognition has the A-I-A pattern over the course of a session, the better his performance. However, inspection of Figure 5 shows that this is largely due to the greater occurrence of the A-I-A pattern with the aesthetics

sessions, which had high accuracy. Although the apparent wide variation of slopes on the different tasks is not statistically significant, it suggests the potential mediating effect that the task can have on the relation between features of cognitive mode and performance accuracy.

Insert Table 9 and Figures 3, 4, and 5 about here

As the engineers gained experience with the task of making formulas, their cognition became more intuitive, it alternated more frequently between analysis and intuition, and their formulas became more accurate (Table 8). Further research will be needed to determine whether these effects are mutually dependent. Their co-occurrence suggests that shifting between analysis and intuition is not just a sign of "floundering about" but is essential to expert reasoning.

Problem 3: The Relation of Alternation Rate to the Judgment Task

The data on the effects of formula task on the alternation rate (Column 2 of Table 8) allow a test of Cognitive Continuum Theory's hypothesis that the rate of shifting between intuition and analysis increases as a function of the stringency of task standards: the aesthetics task should have less rapid alternation than the safety and capacity tasks. The capacity formula task had more rapid alternation between intuition and analysis (-.29) (lower segment to segment correlation, shorter runs) than the aesthetics (.12) and safety (.19) formula tasks. A powerful test of these effects is provided by a repeated

measures analysis of variance of alternation rate as a function of formula task, with ordinal position as a covariate. The data from the 5 subjects for whom all three transcripts were coded are analyzed. The results of this analysis are:

Source	SS	df	MS	F	Sig.
Task	2.320	2	1.160	7.13	.021
Ordinal position (covariate)	1.149	1	1.149	9.13	.019
Within cells	1.139	7	0.163		

The effects of both judgment task and ordinal position on the alternation rate are statistically significant. It was hypothesized that both the safety and capacity tasks would have more stringent standards than the aesthetics task, and hence their cognition would shift between analysis and intuition more frequently. Cognition showed faster alternation on the capacity formula task than on the aesthetics task, as expected, but safety was similar to aesthetics rather than to capacity. This supports Cognitive Continuum Theory's prediction, while casting doubt on the assumption that the safety formula task is perceived by the engineers as having stringent standards. While safety, like aesthetics, lacks a known method of calculation that the engineers could use as a standard, one might assume that due to the high <u>importance</u> of safety the engineers would maintain high standards for themselves on this task, with the resulting frequent alternation between analysis and intuition. If this assumption is mistaken, it would be consistent with other findings from the study of risk--that low probability disasters do not motivate changes in daily behavior (e.g., use of seatbelts).

Problem 4: Dependence of the Mean MBMCCI on Formula Making Task

One of the basic tenets of Cognitive Continuum Theory is that task conditions induce the corresponding mode of cognitive activity. The three formula tasks differ in their "depth task characteristics", as defined and measured by Hammond, Hamm, Grassia, and Pearson (1985): the capacity task is most analysis-inducing (8.22 as measured on the "Task Continuum Index"), the aesthetics task most intuition-inducing (3.15), and safety is in between (6.62) but closer to capacity. The MBMCCI is therefore predicted to be highest for capacity and lowest for aesthetics, with safety in between. This overall pattern was observed (Column 1 of Table 8): aesthetics (MBMCCI = .369), safety (.392), and capacity (.401). The simple correlation between each session's mean MBMCCI and the formula task (measured as aesthetics = 1, safety = 2, and capacity = 3) is .44, p = .039. The more precise repeated measures anova, with ordinal position as a covariate, shows:

Source	SS	df	MS	F	Sig.
Task	.00172	2	.00086	2.11	.192
Ordinal position (covariate)	.00225	1	.00225	5.53	.051
Within cells	.00285	7	.00041		

Although the overall effect of task was not statistically significant ($F(2,7)=2.11$, $p = .19$), the contrast between the aesthetics task and the safety and capacity tasks was ($F(1,2) = 29.47$, $p = .032$). This parallels the finding by Hammond, Hamm, Grassia, and Pearson (1985) that safety was closer to capacity than to aesthetics in its tendency to induce cognitive activity. In addition, it was found that on the average the engineers grew more intuitive when making their later formulas ($F(1,7) = 5.53$, $p = .051$).

Problem 5: The Subtasks and Their Effects on the Engineers' Cognitive Activity

It has been shown that there is moment by moment variation in the position of the engineer's cognitive activity on the Cognitive Continuum, as well as nonlinear trends over time. Further, the topic of the formula and the order in which the engineer produced it have small but significant influences on the mean MBMCCI, and the pace of the alternation is influenced by the formula topic. But no further inquiry into the causes of the moment by moment changes is possible without the more molecular task analysis that Howell (1984) called for.

The coding of the subtask categorization schemes. The categorization schemes pertaining to four informal models of the subtasks the engineer sets for himself were used to investigate whether the observed variation in the MBMCCI is due in part to the self-defined subtask. The mean percent of segments that were assigned to the categories for each of the four independent subtask categorization schemes is shown in Column 3 of Table 10. The reliability of each category's coding is determined with the method used for the categories of the MBMCCI subindices (described above). The reliability correlations for each category (Column 1 of Table 10) range from a low of .05 for the "combine subtask judgments" category of Scheme F to .80 for the "information gathering" category of Scheme II. Percent agreement between the two coders, that is, the percent of two segments on which the two coders gave the same category, was used to produce an overall reliability score for each scheme (Column 2 of Table 10). The Information

Processing Scheme (III) was least reliable, with 56.2% agreement, and the Formula Parts Scheme (I) was most reliable (79.5%). The relative reliability may be viewed as an indication of the appropriateness of the subtask schemes for describing the making of formulas. Finally, an average percent agreement score was calculated for each of the coders responsible for subtask categorization schemes (Table 1).

Insert Table 10 about here

The effects of the subtasks on the MBMCCI. For each scheme, the mean MBMCCI score over all the segments identified as engaged in each subtask can be calculated. Table 10 shows the mean of these scores across sessions for the aesthetics, safety, and capacity formula making sessions separately (Columns 4 to 6) and for all sessions combined (Column 7). The effects of each of the subtask schemes' categories on the engineers' cognition are analyzed with two ANOVA methods that have complementary advantages--one uses data from all sessions, and the other is done for each individual session. The first is a repeated measures anova (3 sessions each for 5 engineers) looking at the mean MBMCCI in each subtask category within each session, controlling for the topic of the formula. Although this "all sessions" analysis addresses all the data in one analysis, it ignores the number of segments that were coded with each category in the individual sessions. For example, the mean MBMCCI score for the 'organizing principle' (Op) subtask (Level I) of a subject who was coded as having the Op category in 5% of 100 segments is given just as much weight, in this

analysis, as the mean MBMCCI for the Op category of a subject who used Op on 20% of 700 segments. Also, since only the mean MBMCCI for each subtask category was used, the "all sessions" analysis tends to underestimate the statistical significance of the results.

The "individual session" analysis is a one-way anova of MBMCCI as a function of the different subtask categories. This analysis takes into account the number of segments that fall into the various subtask categories in the given session. However, due to the non-independence between adjacent segments (discussed above--many of the segments given the same task category were in fact adjacent, so their MBMCCI scores could be dependent), the individual session analyses would overestimate the significance of observed differences in mean MBMCCI between subtask categories.

In each of these analyses, planned orthogonal contrasts were made between subtask categories in addition to an overall test.

Formula Parts (Scheme I). The mean MBMCCI for the Formula Parts categories (Column 7 of Table 10) shows the "whole formula" (mean MBMCCI = .41) and "organizing principle" (.40) categories to be more analytical than the "dimension" category (.37). The differences among these categories are almost significant (in the "all sessions" analysis) at $p = .085$ ($F(2,8) = 3.41$). The contrast of "dimensions" against "organizing principle" and "whole formula" is significant ($F(1,4) = 12.28$, $p = .025$), with no difference between Op and W. The individual anovas reveal that 14 of the 17 sessions had the same pattern (W and Op being more analytical than D), 8

of them significantly so, and that the main exception to the general pattern was Engineer 10's safety formula, where the cognitive activity on the W and D subtask segments had a more analytical mean MBMCCI than on the Op subtask segments. Although it is not surprising that the engineers are more intuitive when attending to the subtask of incorporating individual dimensions into a formula than when thinking about what organizing principle to use, the finding that they use relatively analytical cognition when thinking about the formula as a whole goes against the commonsense idea that thinking about a process as a whole is "wholistic" and hence intuitive.

Search (Scheme II). Due to the small number of categories coded I and R (Column 3 of Table 10), these two subtasks were dropped from this analysis. 'Generation' is the most analytic subtask in this scheme (mean MBMCCI =.42) (Table 10, column 7), 'evaluation' is next (.40), then 'pregeneration' (.37), and finally constraint setting (.35). These differences were highly significant in the "all sessions" analysis ($F(3,12)=10.04$, $p < .001$). While 'generate' and 'evaluate' did not differ significantly ($F(1,4)=.96$), they were more analytical than 'pregenerate' and 'constraint' ($F(1,4) = 32.33$, $p = .005$), and 'pregenerate' was more analytical than 'constraint setting' ($F(1,4) = 7.66$, $p = .05$). Individual session analysis showed the aesthetics session by engineer 12 to be the major exception, with P > C, G, and E. These findings support the idea that the subtasks of a complicated task may vary in the degree to which they are intuitive or analytical, as Payne (1982) suggested.

<u>Information Processing (Scheme III)</u>. Due to insufficient degrees of freedom, it was necessary to analyze this subtask categorization scheme by (a) collapsing the categories into the super categories of control, judgment, and memory, and by (b) comparing only a few categories at a time. Comparison of the control, memory, and judgment supercategories revealed no significant differences ($F(2,8)=1.21$, $p=.348$), though there was a tendency for control to be more analytical than judgment ($F(1,4) = 3.11$, $p = .153$). None of the three control categories (Cg, Cu, Ce) was more analytical than the others. There were nearly significant differences among the memory and judgment categories ($F(3,12) = 3.36$, $p = .055$), with the largest difference being that 'memory storage' is more analytical than 'memory retrieval' and 'mental imagery' ($F(1,4) = 7.02$, $p = .057$). It was surprising that the difference between verbal and numerical judgments in the "all sessions" analysis was not significant ($F(1,4) = 2.44$, $p = .193$), but the result was confirmed by the individual session analyses of variance --for 10 sessions (3 significant) the segments coded 'numerical judgment' were more analytical than the segments coded 'verbal judgment', but the opposite occurred for the remaining 7 sessions (1 significant).

<u>Decomposition (Scheme F)</u>. Structuring the subactivities (MBMCCI = .42) and combining the results of the subactivities (.43) were done with more analytical cognition than naming the whole task's goal (.38), deciding on the breakdown into subactivities (.37), or doing the subactivities (.36). The "all sessions" analysis showed significant differences among the means of these categories ($F(3,12)=7.18$, $p = .005$). Contrasts among the categories showed: 'naming' and 'breaking down' were more intuitive

than 'structuring' ($F(1,4)=7.83$, $p = .049$), and combining subactivity results was more analytical than doing the subactivities ($F(1,4) = 23.18$, $p = .009$). This confirms the finding from the Formula Parts Scheme that the most detailed subtasks are done relatively intuitively. However, the subtask categories are more differentiated here, allowing the "wholistic" approach to the task ('naming, registering, redefining the goal') to be measured as relatively intuitive, as the common sense theory would expect. In this respect, the Decomposition Scheme is better than the Formula Parts Scheme.

In conclusion, it was feasible to measure the engineer's self-defined subtasks, and these subtasks influence the engineers' mode of cognitive activity. The Search and Formula Parts subtask categorization schemes were more reliably coded, indicating that they are more naturally applicable to transcripts of experts producing formulas. The subtasks of the Search and Decomposition Schemes had the greatest influence on the engineers' mode of cognitive activity. In the Search Scheme, the subtasks of generating formula parts and evaluating them were performed with the most analytical cognition, the pregenerational activity was next, and the subtask of setting constraints to guide the subsequent generation of formula parts was done most intuitively. For the Decomposition Scheme, the subtasks that involve detailed coordination of other activities (structuring and combining) used more analytical cognition than the most abstract subtasks (naming the overall goal, breaking it down into subactivities) or the most concrete subtask (performing the subactivities). As for the Formula Parts Scheme, the subtask of dealing with individual dimensions was done more

intuitively than the subtasks of thinking about the formula as a whole or thinking about the organizing principle. There were few differences in mean MBMCCI among the categories in the Information Processing Scheme, and the verbal and numerical judgment processes were done equally analytically.

The Relation of Formula Accuracy to Cognitive Activities and to Subtasks

The detailed "molecular" coding of the engineers' formula making transcripts using categorization schemes pertaining to both the variety of cognitive activities they engaged in and the kinds of subtasks they set for themselves provides an opportunity to determine whether any of these activities are related to the accuracy of the formula the engineer produced. (See Svenson (1985), who has previously related judgment strategies discovered via protocol analysis to performance accuracy.)

Effects of specific MBMCCI subindices. Although achievement is not related to the mean MBMCCI on the session (their correlation is -.19), the MBMCCI index is made up of several subindices. The relation between achievement and the various activities that constitute the subindices of the MBMCCI is captured by the correlation of r_a, over the 17 sessions, with the percent of the categories in which the engineer did the activity (Column 4 of Table 6). (More precise statistical analyses will not be presented here.) It was found that the more the engineer used memory in making judgments (the most intuitive use of memory in Scheme C), the better his achievement. In Scheme D, the more the engineer used pure analysis (calculation, logic, mathematics), the more accurate his formula. The more the engineer thought about the qualities of objects or dimensions (the most

intuitive activity in Scheme G), the better his performance. Thus features of relatively intuitive cognitive activity, as well as features of relatively analytical cognitive activity, are related to better performance in making formulas. However, inspection of Table 6 reveals that in Schemes A, B, C, D, and G, the use of the more analytical cognitive activities, as opposed to the more intuitive ones, tended to be related to the production of less accurate formulas.

Effects of subtask categories. Formula accuracy may be due to the engineer's allocation of effort among the various subtasks of producing the formula. The correlations between achievement and the proportion of time each engineer spent on the various subtasks are shown in Column 8 of Table 10.

In the Formula Parts Scheme, the greater the proportion of the session that the engineer spent focussing on specific dimensions, rather than thinking about the problem of how to produce a formula, the more accurate the resulting formula. In the Search Scheme, the higher the proportion of time the engineer spent prospectively setting constraints on what the formula parts should be, and the lower the proportion of time he spent retrospectively evaluating the formula parts, the better the formula's accuracy. While it has been recognized that trial and error learning is inefficient (Adelman, 1981), this finding suggests that even in the absence of external feedback about the quality of one's actions, it is better to evaluate steps before taking them rather than after. This is significant because the value of setting constraints on possible solutions has been recognized in other contexts (Winston, 1984). Incidentally, working on

dimensions and setting constraints are both subtasks that were performed relatively intuitively (Column 7 of Table 10). Accuracy was not related to the proportion of time the engineer spent on various subtasks of the Decomposition or Information Processing Schemes.

Conclusions. It has proven to be feasible to measure variation in the subject's cognitive activity on a moment by moment basis, through the analysis of verbal protocols. The non-random nature of this moment to moment variation has been demonstrated, and measurement of the rate of alternation on the micro level and of linear and nonlinear trends on the macro level has been accomplished. There is evidence for nonlinear trends, particularly in the aesthetics formula making task, on which cognitive activity started relatively analytically, became more intuitive, and then became analytical again. No overall relation was found between these patterns of cognitive activity over time and the accuracy of the engineer's formula, but the data suggest that the relations differ according to the formula topic. Two predictions from Cognitive Continuum Theory were supported by the data: (a) the average segment was more analytical on the capacity task than on the aesthetics task, and (b) the rate of alternation between analysis and intuition was faster on the capacity task, which has high standards, than on the aesthetics task. The subtasks the engineer set for himself influenced his cognitive mode, particularly the subtasks in the Search and Decomposition categorization schemes. Formula accuracy was dependent both on analytical cognitive activities (such as the use of mathematics and logic) and on intuitive cognitive activities (such as making judgments of qualities). Finally, accuracy was related to attending

to dimensions rather than to whole formulas (in the Formula Parts Scheme) and to constraining the possibilities ahead of time rather than evaluating them afterwards (in the Search Scheme).

Discussion

Verbal protocols were coded on a moment by moment basis in order to (a) register the changes in cognition over time that Polanyi (1958) described, and (b) describe the task on the molecular scale that Howell (1984) suggested is needed in order to advance our understanding of the relation between task features and cognitive mode in complicated tasks.

General Applicability of the Method

Because the Moment by Moment Cognitive Continuum Index is the weighted average of fourteen separate activities that reflect the degree of intuition or analysis in cognition, and because each of these activities is defined in general terms, it is possible to apply this procedure with little change to a broad set of tasks. The techniques were applied in a pilot study to a medical student's discussion with a teaching physician about a case of acute respiratory failure. The relation between the student and physician over time was observed (the teacher was consistently more analytical), as well as the relative use of analysis and intuition in discussion of information from different stages of the case history (the teacher talks about the laboratory reports most analytically, while the student talks about the physical exam most analytically). In addition to the insight gained about the particular situations to which the procedure

is applied, the generality of the procedure allows comparisons across situations by means of such measures as the rate of alternation between analysis and intuition and the linear and nonlinear trend patterns.

The coding of subtasks employed abstract concepts that, except for the Formula Parts Scheme, are not particular to constructing formulas. The scheme for coding subtasks is therefore generally applicable to other unstructured tasks in which the subject needs to produce a complicated model, structure or design. This procedure had some problems that did not arise with the MBMCCI, however. Often several categories in a scheme could be applied to a single segment (e.g., for the Information Processing Scheme, a single segment could involve following a plan, making a judgment, and storing the planned judgment in memory), and so it was necessary to prioritize the categories (see the Coder's Manual). Second, multiple measures of the same concept were not used in the subtask coding schemes, and so they did not have the high reliability that averaging produces. Third, although coders were guided by context, the subtasks were identified only at the level of the individual segment. The coding scheme is therefore not sensitive to the subtask as an event situated in time, with a beginning and an end and a position in a sequence of other subtasks. Despite these comparative disadvantages, the subtasks were sufficiently well defined and coded reliably enough that differences in mean MBMCCI between subtasks could be observed.

Suggestions for Improving the Coding Method

There are several ways to improve the coding procedures with respect to the criteria of reliability, validity, and cost of coding, but they can not be applied simultaneously because the criteria are conflicting. Using additional coding schemes can increase both the reliability and validity of the MBMCCI measures (and the indices of change derived from it), yet only at the cost of additional labor.

Reliability. The MBMCCI is a weighted average of fourteen noisy subindices. Its relatively high reliability (estimated at .78) depends on two factors: (a) the averaging process that lets the signal come through the noise, and (b) the care taken to ensure the reliability of the individual subindices, e.g., through meticulously defining the coding criteria and training the coders to a high standard. Reliability can be increased by either coding more indices or increasing the reliability of the current indices, each of which adds to the costs. Conversely, one can decrease the cost of the coding process only by dropping subindices or decreasing the effort devoted to coding each one reliably.

In this study all of the coded subindices were used in producing the MBMCCI, including those which proved to have low reliability. Perhaps unreliable indices should be dropped in future studies. However, one reason that some subindices were unreliable is that some kinds of behavior addressed by the categories occured only rarely during the task of constructing formulas. The primary coder and the checker had little opportunity to refine the coding of rare behavior, such as the categories in Scheme E. Since these categories occur frequently in medical diagnosis, the task in the pilot study mentioned above, coders would be more reliable

with Scheme E in this application. Therefore, the decision whether to use a coding scheme should be based on its anticipated reliability in the planned application.

Validity. Each of the MBMCCI subindices was selected a priori, on the basis of its theoretical relation to analysis and intuition. Thus the averaging which produced high reliability also produces high face validity. But the convergent validity of the subindices is structurally limited. Because the size of each segment is quite small, its content is limited, and it is not possible for one segment to exhibit the analytical features of many subindices at once. Hence the subindices can not exhibit the high intercorrelation that would signal their convergent validity. Further, since the subindices have low redundancy, it is not possible to drop subindices without threatening to make the operationalization of the cognitive continuum incomplete.

A second validity issue is the low correlation between the mean MBMCCI for a session and the Cognitive Continuum Index (CCI) assigned to the session by Hammond, Hamm, Grassia, and Pearson (1985), which is based largely on statistical descriptions of the answers produced by the formula. Although the purposes of these two measures are quite distinct--the CCI is intended to characterize the overall cognitive mode, and the MBMCCI is intended for comparing cognitive mode between moments--still they would be expected to show some relation. A possible explanation is that the CCI as a measure of cognitive mode is founded on the use of statistical models to describe the judgment policies underlying a set of wholistic judgments. The CCI is appropriate for the film strip and bar graph task conditions in

Hammond, Hamm, Grassia, and Pearson (1985), and for comparing the formula making condition to those tasks. But the CCI is maladapted for measuring cognitive activity during the formula making process, compared to the MBMCCI.

Effort. The coding for this study, involving fourteen judgments of the engineer's cognitive activity during a segment and another four judgments of the subtask he was defining for himself, was costly in terms of time, effort, and coder's patience. Future users of this methodology will want to consider whether the costs can be reduced without compromising the reliability or validity of the MBMCCI or the usefulness of the molecular coding of task. There are a number of avenues for reducing the amount of effort used in this kind of study. For the subindices of the Moment by Moment Cognitive Continuum Index,

1. Discard the less valid subindices. In the author's judgment, the behaviors captured by the subindices measuring difficulty verbalizing and confidence are less pertinent to the position of the cognitive activity on the cognitive continuum than the other elements from Hammond's (1980) lists, and can be excluded.

2. Discard indices which, despite being used frequently, are still unreliable. As discussed above, rarely used subindices may be measured as unreliable due to inadequate communication between coders. However, if the cause of unreliability is lack of conceptual clarity, the coding scheme should be dropped. This decision depends on experience coding the transcripts.

3. Collapse rare categories. If categories in a coding scheme are rarely coded and similar in the degree to which they are analytical or intuitive (e.g., the "comparison", "reasons", and "tradeoff" categories in the Decision Making Scheme), they may be combined. Discovering what is "rare" depends on experience with transcripts of the specific problem being studied.

4. Do not overdiscriminate. In the Decision Making Scheme, the distinction between the "action" and "conscious action" categories was difficult to make and contributed little to the differentiation of analytical and intuitive cognition.

For coding transcripts to identify task on a molecular basis,

1. It may be possible to avoid the need to code subtasks by structuring the task situation. For example, the use of computerized presentation of dynamic decision situations (e.g., Howell, 1984) would define the subtask unambiguously at each moment so that it would not be necessary to code the transcripts in order to identify subtasks.

2. Reduce the scope of the task by defining smaller problems.

3. Even if the researcher is interested in studying cognition on large, unstructured tasks, work can be reduced by selecting just one appropriate subtask coding scheme.

Following these suggestions, the analysis of verbal protocols is a feasible tool for investigation, at the molecular level, of temporal changes in

cognitive activity during judgment, decision making, and problem solving tasks.

References

Adelman, L. (1981). The influence of formal, substantive, and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. Organizational Behavior and Human Performance, 27, 423-442.

Anderson, B. F., Deane, D. H., Hammond, K. R., McClelland, G. H., & Shanteau, J. C. (1981). Concepts in judgment and decision research: Definitions, sources, interrelations, comments. New York: Praeger.

Behn, R. D., & Vaupel, J. W. (1982). Quick analysis for busy decision makers. New York: Basic.

Bradshaw, G. L., Langley, P. W., & Simon, H. A. (1983). Studying scientific discovery by computer simulation. Pittsburgh: Carnegie-Mellon University, Department of Psychology.

Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.). Berkeley: University of California Press.

Buchanan, B. G., & Shortliffe, E. H. (1984). Knowledge engineering. In B. G. Buchanan & E. H. Shortliffe, (Eds.), Rule based systems: The MYCIN experiments of the Stanford Heuristic Programming Project (pp. 149-158). Reading, Mass.: Addison Wesley.

Davis, R. (1979). Interactive transfer of expertise. Artificial Intelligence, 12, 121-157.

Edwards, W. & Newman, J. R. (1982). Multiattribute evaluation. Beverly Hills, CA.: Sage Publications.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, Mass.: The MIT Press.

Fogiel, M., & Staff of Research and Education Association. (1978). The statistics problem solver. New York: Research and Education Association.

Goldsberry, B. S. (1983). In search of the components of task induced judgment decrements (Technical Rep. No. 83-3). Houston, TX: Rice University.

Hamm, R. M. (1985). Coder's manual. Codebook for categorization of highway formula protocols. Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1980). The integration of research in judgment and decision theory (Report No. 226). Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R. (1981). Principles of organization in intuitive and analytical cognition (Report No. 231). Boulder, Colorado: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R., Anderson, B. F., Sutherland, J., & Marvin, B. (1984). Improving scientists' judgments of risk. Risk Analysis, 4, 69-78.

Hammond, K. R., Hamm, R. M., & Grassia, J. (1985). Generalizing over conditions by combining the multitrait multimethod matrix and the representative design of experiments. Boulder, CO: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R., Hamm, R. M. Grassia, J., & Pearson, T. (1985). Direct comparison of the relative efficacy of intuitive and analytical cognition. Boulder, Colorado: University of Colorado, Center for Research on Judgment and Policy.

Holling, C. S. (Ed.). (1978). Adaptive environmental assessment and management, New York: John Wiley and Sons.

Howell, W. C. (1984). Task influences in the analytic-intuitive approach to decision making. Houston, TX: Rice University, Department of Psychology.

Isenberg, D. J. (1984). How senior managers think. Harvard Business Review, 62, 81-90.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. Econometrica, 47, 263-291.

Keeney, R. L., & Raiffa, H. (1976). Decisions with multiple objectives: Preferences and value tradeoffs. New York: John Wiley & Sons.

Kirwan, J. R., Chaput de Saintonge, D. M., Joyce, C. R. B., Holmes, J., & Currey, H. L. F. (1985). Inability of rheumatologists to describe their true policies for assessing rheumatoid arthritis. Manuscript submitted for publication.

Lindsay, P. H., & Norman, D. A. (1972). Human information processing: An introduction to psychology. New York: Academic Press.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Payne, J. W. (1982). Contingent decision behavior. Psychological Bulletin, 92, 382-402.

Pepper, S. (1948). World hypotheses. Berkeley: University of California Press.

Polanyi, M. (1958). Personal knowledge: Towards a post-critical philosophy. New York: Harper & Row.

Schwartz, D. R., & Howell, W. C. (in press). Optional stopping performance under graphic and numeric CRT formatting. Human Factors.

Stein, M. I. (1974). Stimulating creativity. New York: Academic Press.

Svenson, O. (1985). Cognitive strategies in a complex judgment task: Analyses of concurrent verbal reports and judgments of cumulated risk over different exposure times. Organizational Behavior and Human Decision Processes, 36, 1-15.

Westcott, M. R. (1967). _Toward a contemporary psychology of intuition_. New York: Holt, Rinehart & Winston.

Winston, P. H. (1984). _Artificial intelligence_ (2nd ed.). Reading, Mass.: Addison Wesley.

Appendix A

Formula-production Instructions

HIGHWAY JUDGMENT PROJECT 1982 **A.E.INST**

NAME: ______________________ **DATE:** ______________________

FORMULA PRESENTATION.
INSTRUCTIONS.

AESTHETIC VALUE JUDGMENT TASK.

In this session your task is to develop a general procedure for making a judgment of the aesthetic value of highways. That is, you will make up a formal procedure, such as a mathematical equation, that will take a highway, described in terms of its important characteristics, and produce a number that reflects its aesthetic value as accurately as possible. You may use any kind of procedure for translating from the numerical highway description to the numerical judgment, as long as it produces a single number, in the appropriate units, representing the aesthetic value of the highway. In your final formula, then, all constants must be numbers.

Your formula must be applicable to two-lane rural Colorado highways, similar to those you have judged in the earlier sessions of this study. Thus, you should be thinking about already existing highways, rather than the designs of new ones that would be built to current standards.

In making your formula for aesthetic value, you will have available the definition of aesthetic value and a list of some of the characteristics of highways which may be important to your judgment. You may use any of these characteristics in composing your formula, although you do not have to use them all. But do not use any characteristics that do not appear on the list.

The units in which the characteristics are listed are given on the dimension definition sheet. The ranges of possible values over which the characteristics of the highways may be expected to vary are also listed. Your formula need not handle any highways with values outside these ranges.

Your task, then, is to construct a formal procedure, using those highway characteristics that you think are appropriate, that will estimate as accurately as possible the aesthetic value of highways. You should spend approximately as much time on this task as you did making judgments about film strips or bar graphs. Of course, we are not giving you time to research the whole problem--just use your current knowledge--but we would like to see what you can do when you are using your best analytical judgment.

We are providing some materials for you to work with: scratch paper, graph paper, pencils, a ruler, and a calculator. Please label and save your scratch sheets to give to the researcher when you have finished.

When you are satisfied with your formula and are ready to write it out completely, please tell the researcher.

One final note: please do not discuss this session with any participating engineers until everyone has completed this last phase of our study.

Appendix B

Think-aloud Instructions

HIGHWAY JUDGMENT PROJECT 1982 A.THINK

NAME: ______________________ DATE: ______________________

FORMULA PRESENTATION.
THINK ALOUD INSTRUCTIONS.

(Researcher reads this to engineer:)

In this session I am going to ask you to produce a formal procedure for expressing your judgment, and to THINK ALOUD as you work on it. What I mean by think aloud is that I want you to tell me EVERYTHING you are thinking from the time you are first given the next sheet of instructions, until you are satisfied with the formula you have produced. This includes reporting visual images, metaphors, or schemes you might think of, even though they are not present as words in your head. I would like you to talk aloud constantly from the time that you start reading the instructions until you have finished the formula, including saying what you are reading and what you are writing. I don't want you to try to plan out what you are going to say or to try to explain what you are saying to me. Just act as if you are alone in the room speaking to yourself. I'll be taperecording what you say. When you hear a high-pitched signal from my tape recorder, please pause while I change the cassette.

It is most important that you keep talking. If you are silent for any long period of time I will remind you of this. In order to monitor you work without otherwise intruding, I will sit behind you. Do you have any questions?

For practice, now I want you to think aloud as you answer the question,

"How many windows are there in your house?"

Any questions?

You will start by reading aloud these definition sheets.

Appendix C

Definition of Aesthetic Value

HIGHWAY JUDGMENT PROJECT 1982 IQA.E.INST.DEF

NAME: ______________________ DATE: ______________________

INTRODUCTION. AESTHETIC VALUE JUDGMENT TASK.

The aesthetic value of a highway involves its beauty; how pleasing it is to see and to use. As you know, the aesthetic value of highways is important to citizens; they consider it when deciding what route to take on a trip, where to live, or where to locate a resort. They tend to evaluate the aesthetic value of highways by taking into account the whole scene -- the roadway itself, the surrounding land, and the relation between the two.

The aesthetic value that citizens place on highways is important to highway engineers as they plan new highways and maintain or rebuild old ones. Consequently, it is important for highway engineers to be able to judge or predict how pleasing a highway will appear to citizens.

We will be asking you to predict average citizens' aesthetic evaluations of a number of 2-lane rural highways, using a scale that runs from 1 to 10, where 1 means that the citizens would think that the highway has very low aesthetic value, and 10 means that they would think it has very high aesthetic value. Values of 5 or 6 mean the citizens think the highway is roughly halfway between the best and the worst.

Appendix D

Definition of the Dimensions to be Used in the Formula

HIGHWAY JUDGMENT PROJECT 1982 **E.DIM.DEF**

NAME: ____________________ **DATE:** ____________________

DEFINITIONS OF HIGHWAY DIMENSIONS FOR AESTHETIC VALUE JUDGMENT TASK.

1. Attractiveness of design. This indicates how attractive the design of the highway is, considering the materials and form of the road surface, bridges, entrances, curves, etc. The scale goes from 1=very unattractive to 10=very attractive.

2. Road Condition. This indicates whether the road is decayed, in poor repair, showing signs of neglect, or is on the other hand in good repair, for example with smooth surfaces and clear markings. The scale ranges from 1=in very bad condition to 10=in very good condition.

3. Scenery. This indicates whether the natural landscape surrounding a road, all the way out to the horizon, is dull or magnificent. The scale goes from 1=very dull scenery to 10=very beautiful scenery.

4. Roadside culture and artifacts. This measure indicates the amount and attractiveness of the human artifacts that may be seen around the highway, including the homes, buildings, utility poles, intersections, railroad crossings, stores, farms, parking lots, factories, mines, advertisements, trash, etc. The scale goes from 1=negligible where except for the road itself there is no sign of human culture, to 10=overwhelming, where the are human artifacts in every direction.

HIGHWAY JUDGMENT PROJECT 1982 **E.DIM.DEF**

NAME: ______________________ **DATE:** ______________________

DEFINITIONS OF HIGHWAY DIMENSIONS FOR AESTHETIC VALUE JUDGMENT TASK. PAGE 2

5. Landscaping. This measure indicates the overall quality of the landscaping and repair that have been done on the land around the highway. Landscaping both fixes damage that was done in making the road, and is attractive in and of itself. Both considerations are included in this overall evaluation of the quality of the landscaping. The scale goes from 1=very poor landscaping to 10=very good landscaping.

6. Color. This measure indicates the quality of the color and lighting one experiences as one drives on the particular highway. It can range from dull, monotonous and grey to bright, dramatic and colorful. The scale goes from 1=very dull to 10=very colorful.

7. Vegetation. This measure indicates the amount of vegetation in the landscape, including lichen and moss, grass and weeds, and shrubs and trees. The scale ranges from 1=no vegetation at all to 10=very lush vegetation.

8. Terrain. This measure indicates whether the terrain surrounding the highway is plains, rolling, or mountainous, as defined and measured by the Colorado Department of Highways. The scale is 1=plains, 2=rolling, and 3=mountainous.

Author Notes

Thanks are due to Carole Chrvala, Janet Grassia, and Rich Ling for diligent coding and to Ken Hammond and Janet Grassia for comments on earlier drafts of the paper.

Footnotes

[1]The detailed definition of each category may be found in a separate Coder's Manual (Hamm, 1985) available from the Center for Research on Judgment and Policy. For ease of reference, the letter codes from the Coder's Manual are used here.

[2]The expected number of runs is 2*k*(N-K)/N (Fogiel and Staff, 1978, p. 932). The expected run length is therefore N/(2*K*(N-K)/N) or (N**2)/2*K*(N-K). If K = N-K = N/2, this becomes ((2K)**2)/2*K*K or 2.

Table 1

Reliabilities of Categorization Schemes

Levels	Coded by	Sessions Coded	Checked by	Reliability: Correlation	Reliability: Percent Agreement
I,II,III,IV	RL	10c,12e,15s	RH		68.8%
A,B,F	CC	1c,5e,8s	RH	.42 (for A,B)	67.4% (for F)
C,D,E,G,H	RH	8s,10c,12e	JG	.52	
J,K	JG	5s,8e,10c	RH	.69	

Table 2

Cognitive Activity Categorization Scheme, with Scores for Calculating the MBMCCI Subindices

Score	Scheme A. Decisions and decision making (dm).
0	Act - action without justification
1	Cs - conscious of taking action
2	Opt - describes options
3	Ev - evaluate or negate an action or option
4	Com - dm involving comparison without discussion
5	Rea - dm involving simple reasons and justifications
6	Tra - dm involving tradeoffs among dimensions

Score	Scheme B. Justifications.
0	No jus (ir) - no justification
1	Jus - use of any form of justification

Score	Scheme C. Kind of Memory.
0	Jd - use of memory in making a judgment
1	Ep - episodic memory
2	Se - semantic memory

Score	Scheme D. Source of Knowledge Used
0	Ex - Experience
1	Ap - Applied science, engineering
2	Ba - Basic science
3	Pu - Pure analysis: mathematics and logic

Score	Scheme E. Knowledge about Co-occurrences: Causality and Correlation
0	Ir - no co-occurrence knowledge was used
1	Cor - correlation, non-causal
2	Amb - ambiguous causal relation
3	Pre - prediction, causal
3	Dia - diagnosis, causal

Score	Scheme G. Qualities and relations.
0	Qua - stating judgments of qualities
1	Com - making comparisons
2	Rel - stating relationships

Score	Scheme H. Quantities, numbers.
0	Vag - using vague quantitative terms
1	Rat - using numbers as ratings
2	Var - using variables and formulas

Scheme J. Difficulty verbalizing.*
- Pa - pauses and hesitations
- Re - rephrasings and repetitions
- Ch - changing sentence structure in midstream
- Inc - incomplete sentences
- Mut - muttering and inaudible verbalization

Scheme K. Confidence.*
- Dbt - expressions of doubt
- Cnf - expressions of confidence

* Scores for Schemes J and K consisted of the count of incidents of the category in each segment.

Table 3

MBMCCI Subindex Category Codes and Calculated MBMCCI for Segments in Figure 2

SEG	MBMCCI	A	B	C	D	E	G	H	Scheme J Pa	Scheme J Mut	Scheme J Re	Scheme J Ch	Scheme J Inc	Scheme K Dbt	Scheme K Cnf
258	.45	act	ir	jd	ex	pre	rel	ir	0	0	0	0	0	0	0
259	.36	act	ir	jd	ex	pre	rel	vag	0	1	0	0	1	0	0
262	.46	act	ir	jd	pu	ir	ir	ir	0	0	1	0	0	0	0
263	.33	act	ir	se	ap	ir	qua	vag	0	0	0	0	0	0	0
264	.48	act	ir	se	ap	pre	rel	vag	0	0	1	0	0	0	0
265	.52	cs	ir	se	ap	pre	rel	vag	1	0	0	0	0	0	0
266	.40	cs	ir	ir	ir	ir	ir	ir	0	0	0	0	0	0	0
268	.40	cs	ir	ir	ir	ir	ir	ir	0	0	0	0	0	0	0
269	.35	cs	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
270	.32	act	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
271	.31	cs	ir	jd	ex	ir	ir	ir	0	0	0	0	0	0	0
272	.41	act	ir	jd	ex	ir	rel	rat	0	0	0	0	0	0	0
273	.30	cs	ir	jd	ex	ir	qua	rat	0	2	0	0	0	0	0
274	.32	act	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
275	.35	act	ir	ir	ir	ir	qua	rat	0	0	1	0	0	0	0
276	.41	act	ir	jd	ex	pre	rel	vag	0	0	0	0	0	0	0
277	.32	act	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
278	.29	act	ir	jd	ex	ir	qua	rat	0	0	1	0	0	0	0
279	.39	act	ir	jd	ex	pre	rel	vag	0	1	0	0	0	0	0
280	.51	cs	ir	jd	ex	pre	rel	rat	0	1	0	0	0	0	0
281	.53	cs	ir	jd	pu	ir	qua	rat	0	0	0	0	0	0	0
282	.53	cs	ir	jd	ex	pre	rel	rat	0	0	0	0	0	0	0
283	.35	cs	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
284	.32	act	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
285	.51	cs	ir	jd	ex	pre	rel	rat	0	1	0	0	0	0	0
286	.32	act	ir	jd	ex	ir	qua	rat	0	0	0	0	0	0	0
287	.30	cs	ir	jd	ex	ir	qua	rat	1	1	0	0	0	0	0
288	.43	cs	ir	jd	ex	pre	rel	ir	0	1	1	0	0	0	0
289	.39	cs	ir	jd	ex	pre	rel	vag	0	1	1	0	0	0	0
290	.42	cs	ir	jd	ex	pre	rel	vag	0	1	0	0	0	0	0
291	.53	cs	ir	jd	ex	pre	rel	rat	0	0	0	0	0	0	0
292	.37	cs	ir	ir	ir	ir	ir	ir	0	1	0	0	0	0	0
294	.53	cs	ir	jd	ex	pre	rel	rat	0	0	0	0	0	0	0
295	.44	cs	ir	jd	ex	pre	rel	vag	0	0	0	0	0	0	0

Table 4

Subtask Categorization Schemes

Scheme F. Stages of Analytic Decomposition Process.	
Na	Name, register, redefine the goal.
Br	Break judgment task into smaller tasks.
St	Establish structure for relating subtask results.
Ju	Make the subtask judgment and state it, remember it.
Co	Combine subtask judgments.
Scheme I. Judgment Analysis.	
W	Whole formula
OP	Organizing Principle
D	Dimension
Scheme II. Search.	
I	Information gathering
P	Pregenerational activity, familiarization
C	Constraint setting, focussing
G	Generate a formula or formula part
E	Evaluate or justify a formula or formula part
R	Report a formula or formula part
Scheme III. Information Processing.	
	Control
Cg	Generate plans, goals, or procedures
Cu	Use plans
Ce	Evaluate plans or their results
	Memory
Ms	Store in memory
Mr	Retrieve from memory
Mi	Imagine
	Judgment
Jv	Verbal judgment
Jn	Numerical judgment
Scheme IV. Type of Verbalization	
Con	Concurrent verbalization
Ret	Retrospective report

Table 5

Subtask Category Codes for Segments in Figure 2

SEGMENT NUMBER	LEVI	LEVII	LEVIII	LEVF
257	d	p	mr	ju
258	d	p	mr	ju
261	op	p	cg	br
262	op	p	cg	br
263	d	p	mr	ju
264	d	p	mr	ju
265	d	p	cu	ju
267	d	g	cg	ju
268	d	p	mr	ju
269	d	p	cu	ju
270	d	p	cu	ju
271	d	p	cu	ju
272	d	g	cu	ju
273	d	p	cu	ju
274	d	p	cu	ju
275	d	p	cu	ju
276	d	p	cu	ju
277	d	p	cu	ju
278	d	p	cu	ju
279	d	p	cu	ju
280	d	p	cu	ju
281	d	g	cu	ju
282	d	p	cu	ju
283	d	p	cu	ju
284	d	p	cu	ju
285	d	g	cu	ju
286	d	p	cu	ju
287	d	p	cu	ju
288	d	p	cu	ju
289	d	p	cu	ju
290	d	p	cu	ju
291	d	g	cu	ju
293	d	p	cu	ju
294	d	g	cu	ju
295	d	g	jn	ju

Table 6

MBMCCI Subindex Categorization Schemes

	Reliability[a] Category	Reliability[a] Scheme	Mean Percent[b]	Correlation % with r_a[c]
Scheme A. Decisions		.26		
Act - action	.40		56.2	.23
Cs - consciousness	.38		34.6	-.01
Opt - options	.29		1.8	-.07
Ev - evaluate	.27		3.0	-.29
Com - comparison	-.03		0.3	-.35+
Rea - reasons	.36		3.3	-.22
Tra - tradeoffs	~		0.02	-.15
Scheme B. Justifications.		.36		
Jus - justifications	.36		6.4	-.30
Scheme C. Memory.		.42		
Jd - judgment	.31		32.3	.45*
Ep - episodic	.58		0.6	-.29
Se - semantic	.38		8.2	-.16
Scheme D. Knowledge		.43		
Ex - Experience	.58		32.1	.33+
Ap - Applied science	~		1.8	.12
Ba - Basic science	~		0.0	~
Pu - Pure analysis	.31		7.3	.43*
Scheme E. Causality		-.02		
Cor - correlation	~		0.8	.34+
Amb - ambiguous	~		0.2	.11
Pre - prediction	.77		4.3	.04
Dia - diagnosis	~		0.1	-.53*
Scheme G. Qualities.		.53		
Qua - qualities	.51		24.2	.42*
Com - comparisons	.70		6.2	-.06
Rel - relations	.48		18.5	-.13
Scheme H. Quantification.		.86		
Vag - vague	.63		13.4	.19
Rat - ratings	.76		26.1	.13
Var - variables	~		2.3	.19
Scheme J. Trouble speaking		.81		
Pa - pauses	.97			
Re - rephrasings	.76			
Ch - changes	.13			
Inc - incomplete	.32			
Mut - muttering	.77			
Scheme K. Confidence.		.50		
Dbt - doubt	.35			
Cnf - confidence	.71			

[a]Coding reliabilities for categories and whole subindices.
[b]Mean percents of segments coded in each category.
[c]Correlation between achievement and percent of segments coded in each category.

Table 7

Intercorrelations Among Elements of the Alternation Rate Index

	Autocorrelation Analysis					Runs Analysis	
	Lag 1	Lag 2	Lag 3	Lag 4	# Pos	-Z	Runlength
Stability Index	.71	.72	.81	.78	.83	.84	.88
Lag 1 autocorrelation		.73	.76	.60	.63	.46	.49
Lag 2 autocorrelation			.79	.78	.55	.37	.65
Lag 3 autocorrelation				.89	.69	.51	.61
Lag 4 autocorrelation					.68	.46	.61
Number positive ac's						.57	.55
Z-score (reversed)							.77

Note. N = 17.

Table 8

Mean Session Score on Each of Five Indices

		N	Mean MBMCCI	Alternation Rate	Linear Trend	Nonlinear Trend	Achievement (r_a)
	Aesthetics	5	.369	.124	.021	-.188	.923
	sd		(.024)	(.953)	(.031)	(.147)	(.016)
Task	Safety	6	.392	.188	.028	-.001	.348
	sd		(.028)	(.953)	(.076)	(.005)	(.337)
	Capacity	6	.401	-.292	.000	.017	.361
	sd		(.030)	(.887)	(.002)	(.042)	(.539)
oneway	% variance		20.4	7.3	6.8	58.7	36.3
anova	p value		.204	.588	.610	.002**	.043*
repeated	% variance		37.6	66.9	21.4	61.4	36.0
measures	p value		.192	.021*	.430	.036*	.209
	First	5	.405	.252	.048	-.045	.377
Ordinal	sd		(.036)	(1.159)	(.082)	(.104)	(.457)
	Second	6	.388	-.099	.008	-.010	.461
Position	sd		(.023)	(.601)	(.011)	(.084)	(.564)
	Third	6	.375	-.110	-.003	-.093	.705
	sd		(.026)	(.816)	(.004)	(.163)	(.290)
oneway	% variance		18.6	4.1	22.0	8.9	10.4
anova	p value		.238	.747	.176	.520	.465
repeated	% variance		44.1	56.6	34.6	2.1	16.8
measures	p value		.051+	.019*	.096+	.710	.273
Subject	1	2	.392	.718	.000	.004	.214
	sd		(.013)	(.614)	(.000)	(.005)	(.072)
	5	3	.399	-.383	.001	.002	.780
	sd		(.037)	(.242)	(.010)	(.010)	(.138)
	8	3	.354	-.351	.024	-.079	.363
	sd		(.008)	(.967)	(.041)	(.131)	(.818)
	10	3	.411	1.026	.062	-.102	.534
	sd		(.017)	(.877)	(.104)	(.267)	(.515)
	12	3	.393	-.610	-.002	-.047	.572
	sd		(.044)	(.300)	(.004)	(.081)	(.356)
	15	3	.382	-.160	.006	-.058	.567
	sd		(.014)	(.586)	(.020)	(.089)	(.515)
oneway	% variance		41.7	56.5	25.8	10.6	15.1
anova	p value		.246	.068+	.593	.925	.845
All categories			.388	.000	.016	-.050	.522
	sd		(.029)	(.852)	(.049)	(.137)	(.495)

Note. The mean session scores are broken down by formula task, ordinal position, and subject; and results of statistical tests by one-way and repeated measures analysis of variance.

**is $p < .01$. * is $p < .05$. + is $p < .10$.

Table 9

Correlations and Slopes for the Linear Relations Between Achievement and Indices of MBMCCI Change Patterns

Formula Task	Alternation rate		Linear Trend		Nonlinear Trend		Mean MBMCCI	
	r	slope	r	slope	r	slope	r	slope
Aesthetics	-.46	-.01	-.18	-.09	-.12	-.01	-.77+	-.52
Safety	-.72+	-.25	-.55	-2.46	.01	.55	-.48	-5.82
Capacity	.19	.11	.23	52.29	.41	5.30	.46	8.30
Total set	-.06	-.03	-.17	-1.60	-.41	-1.51	-.19	-2.92

+ = $p < .10$.

Table 10

Subtask Categorization Schemes

		Reliability[a] Categ	Reliability[a] Scheme	Mean % in Category[b]	Mean MBMCCI score[c] Aes	Saf	Cap	Mean	Correlation of % with r_a[d]
Scheme F. Decomposition			67.4%	100.0	.37	.39	.40	.39	
Na	Name goal	.17		11.5	.39	.36	.37	.38	.05
Br	Breakdown	.31		3.5	.35	.38	.37	.37	.30
St	Structure	.39		16.4	.41	.39	.44	.42	-.15
Ju	Subactivity	.62		56.8	.32	.38	.39	.36	.13
Co	Combine	.05		11.8	.45	.42	.41	.43	-.24
Scheme I. Formula parts			79.5%	100.0	.37	.39	.40	.39	
W	Whole formula	.72	.	38.4	.40	.41	.42	.41	-.53*
Op	Organizing Princ	.55		7.9	.39	.40	.42	.40	.32
D	Dimension	.73		53.8	.34	.38	.39	.37	.44*
Scheme II. Search			70.6%	100.0	.37	.39	.40	.39	
I	Gather information	.80		14.5	.39	.50	.30	.40	.35+
P	Pregenerate	.51		33.3	.37	.37	.38	.37	.09
C	Constraint setting	.52		13.8	.31	.36	.40	.36	.63**
G	Generate formula	.65		18.7	.40	.43	.42	.42	-.16
E	Evaluate formula	.64		17.4	.35	.41	.43	.40	-.68**
R	Report formula	.79		2.3	.40	.39	.40	.40	.17
Scheme III. Information Processing			56.2%	100.0	.37	.39	.40	.39	
	Control			52.7	.38	.39	.40	.39	
Cg	Generate plans	.53		13.3	.39	.40	.41	.40	.14
Cu	Use plans	.28		35.8	.38	.39	.39	.39	-.33+
Ce	Evaluate results	.56		3.6	.36	.40	.41	.39	-.17
	Memory			25.9	.36	.40	.38	.38	
Ms	Store	.67		15.6	.40	.50	.38	.43	.35+
Mr	Retrieve	.18		9.3	.35	.38	.39	.37	.01
Mi	Imagine	.56		1.0	.33	.36	.40	.36	-.13
	Judgment			21.4	.34	.39	.40	.38	
Jv	Verbal	.67		11.3	.31	.36	.41	.36	.00
Jn	Numerical	.35		10.1	.36	.40	.40	.39	-.02

[a]Coding reliabilities for each category and for whole scheme.
[b]Percent segments coded in each category.
[c]Mean MBMCCI in each category.
[d]Correlation between achievment and percent of segments coded in each category.
** is $p < .01$. * is $p < .05$. + is $p < .10$. N = 17.

Figure Captions

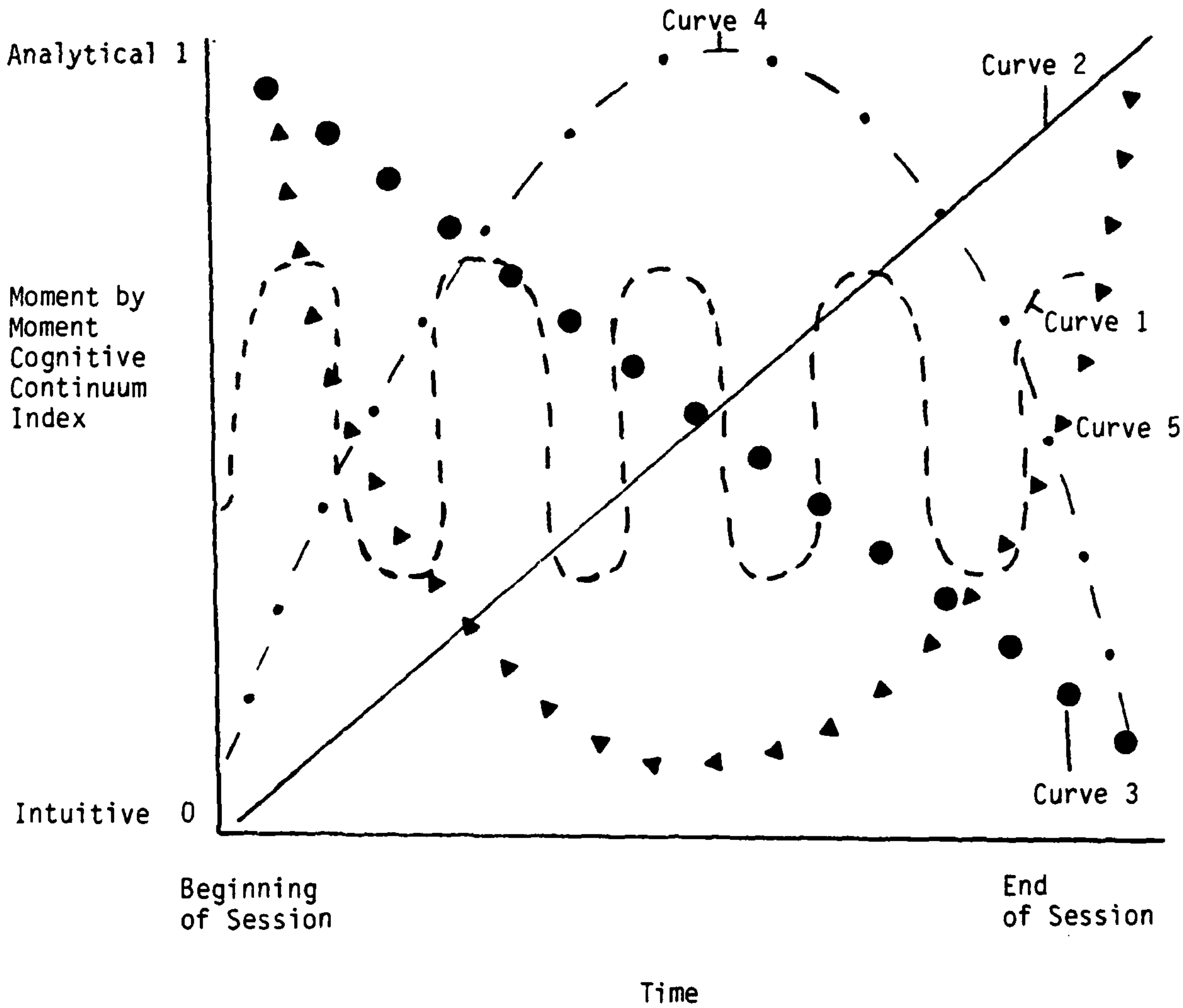
Figure 1. Possible patterns for change in the MBMCCI over time.

Figure 2. Example of segmented and numbered transcript.

Figure 3. Graph of achievement as a function of alternation rate.

Figure 4. Graph of achievement as a function of linear trend.

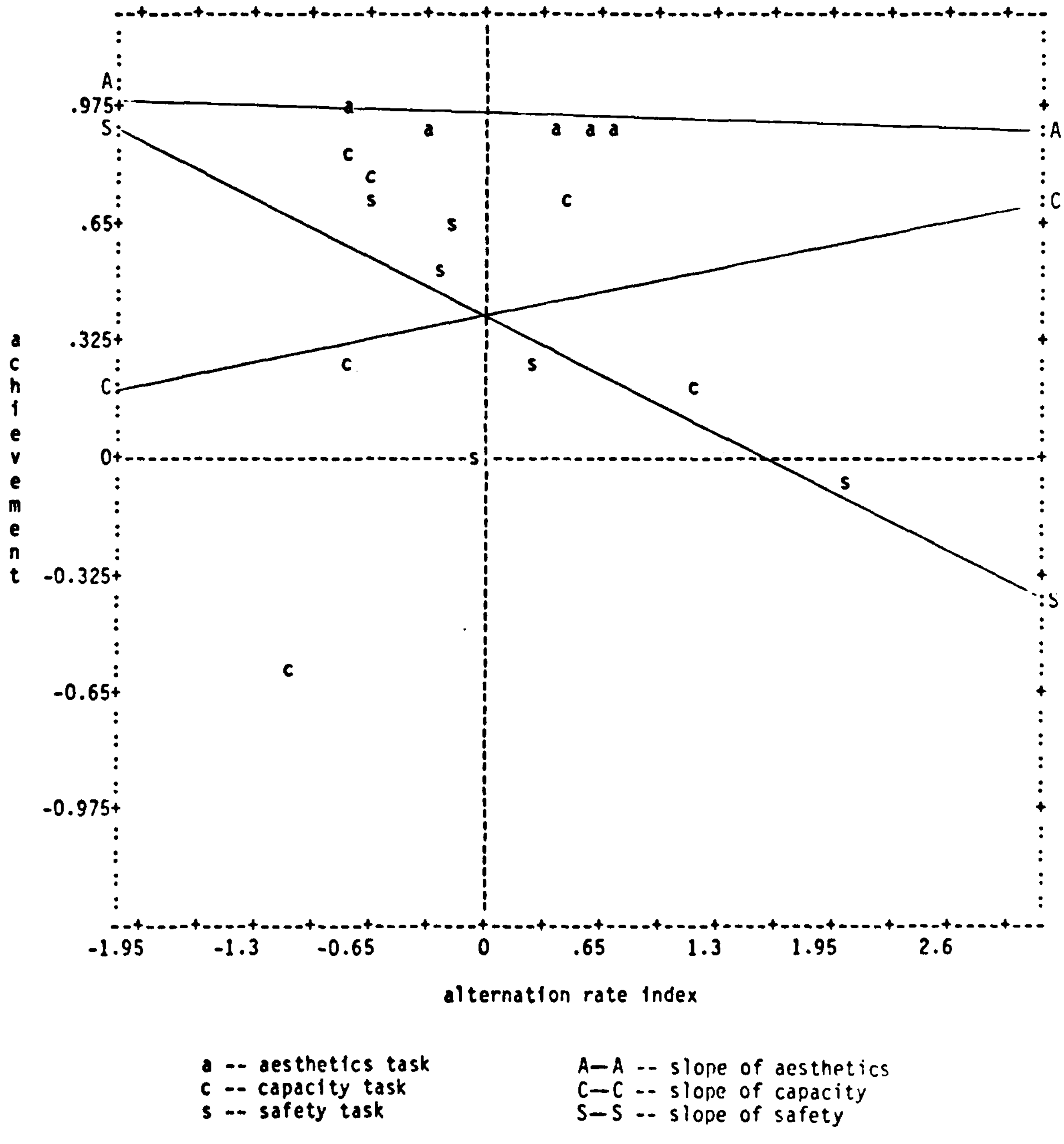
Figure 5. Graph of achievement as a function of nonlinear trend pattern.



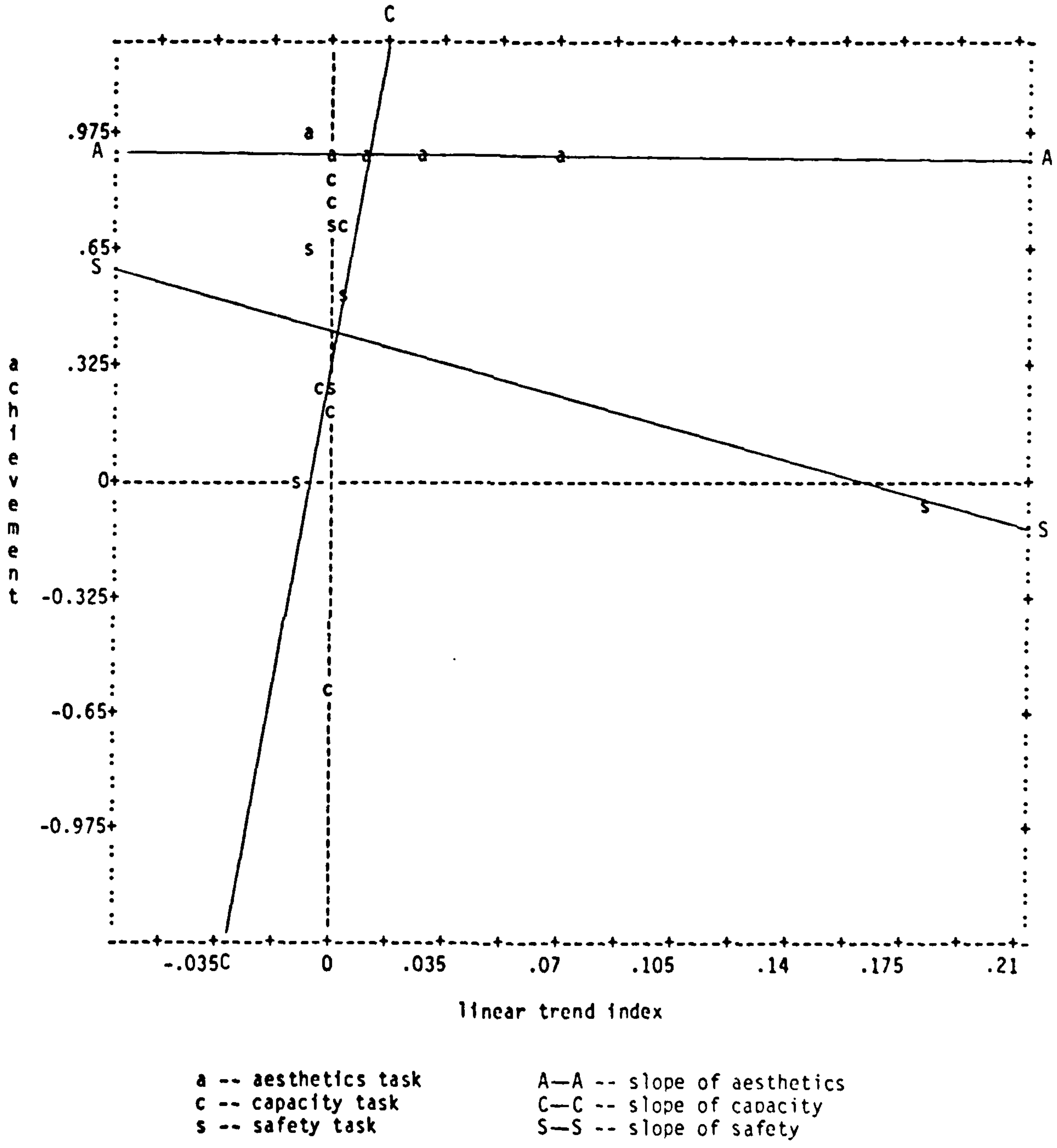
Curve 4
Curve 2
Analytical 1
Moment by
Moment
Cognitive
Continuum
Index
Curve 1
Curve 5
Curve 3
Intuitive 0
Beginning
of Session
End
of Session
Time

[[256]Well, we know that] [[257]if you have very few objects,] [[258]and you have a shoulder width of 7 to 10,] [[259]you're really not going to get,] [[260]well, as a--as a matter of fact,] [[261]what we have to do is we have to create some sort of average condition] [[262]for all--all state highways.] [[263]So considering that many state highways do not have adequate shoulders,] [[264]that--and if we are able to reduce the number of objects and increase the shoulder width,] [[265]then we should have a little bit of PAUSE betterment in the accident rate.] [[266]So that means that our,] [[267]let's say our average condition,] [[268]and I'll denote that by a star,] [[269]would be 5 to 10 objects per mile] [[270]and 3 to 6 feet of shoulder width.] [[271]So that means if we're average] [[272]then our factor of 1.0 goes there.] [[273]And I'd also say that if you have 7 to 10 feet of shoulder width,] [[274]even with 10 to 18 type of,] [[275]10 to 18 objects per mile,] [[276]you're still going to have close to no change.] [[277]As you get 7 to 10 feet with 5,] [[278]with still only 5 to 10,] [[279]you can--you can get somewhere upwards of about--,] [[280]oh, perhaps a 5 to 10 percent reduction in your accident rate.] [[281]Let's put due to a 10 percent reduction,] [[282]the rate would only be 90 percent of normal.] [[283]And let's say with 7 to 10 feet,] [[284]0 to 5,] [[285]I'd say go to .85 of normal.] [[286]And with 10 to 18 objects per mile,] [[287]PAUSE 0 sight distance, [5 syllables missing]] [[288]0, 0 shoulder width, it would create a condition that was--] [[289]it would severely aggrevate the opportunity for accidents] [[290]and could increase your accident rate] [[291]probably by 50 percent.] [[292]In other words,] [[293]that's probably the single most important factor there.] [[294]So I'm going to put 1.5 for that condition] [[295]that would increase the accident rate.]

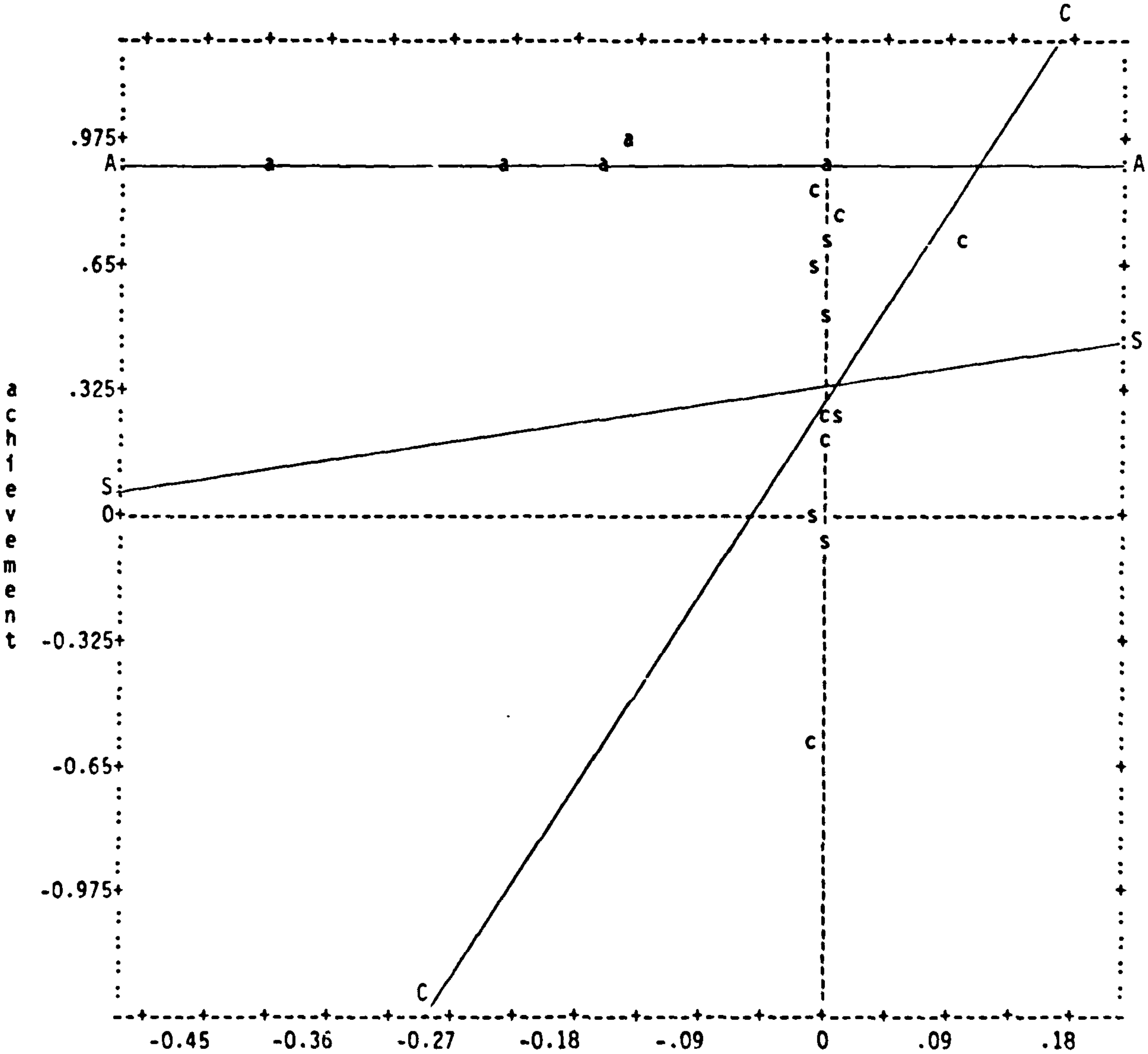
Achievement and Alternation Rate Index



Achievement and Linear Trend Index



Achievement and Nonlinear Trend Index



a -- aesthetics task
c -- capacity task
s -- safety task

A—A -- slope of aesthetics
C—C -- slope of capacity
S—S -- slope of safety

OFFICE OF NAVAL RESEARCH

Engineering Psychology Program

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CAPT Paul R. Chatelier
Office of the Deputy U er Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D. C. 20301

Department of the Navy
Engineering Psychology Program
Office of Naval Research
Code 442EP
800 North Quincy Street
Arlington, VA 22217-5000 (3 copies)

Aviation & Aerospace Technology
Programs
Code 210
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR. Paul E. Girard
Code 252
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Physiology Neurobiology Program
Office of Naval Research
Code 441NP
800 North Quincy Street
Arlington, VA 22217-5000

CAPT. P. M. Curran
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mathematics Program
Code 411MA
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Douglas DePriest
Code 411
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

Information Sciences Division
Code 433
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR Kent S. Hull
Helicopter/VTOL Human Factors
Office
MS 239-21
NASA/Ames Research Center
Moffett Field, CA 94035

Dr. Carl E. England
Naval Health Research Center
Environmental Physiology
P. O. Box 85122
San Diego, CA 92138

CAPT William M. Houk
Commanding Officer
Naval Medical R&D Command
Bethesda, MD 20814-5055

Department of the Navy

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mr. R. Lawson
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

CDR James Offutt
Office of the Secretary of Defense
Strategic Defense Initiative Organization
Washington, D.C. 20301-7100

Director
Technical Information Division
Code 2627
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D.C. 23075-5000

Dr. J. S. Lawson, Jr.
4773-C Kahala Avenue
Honolulu, HI 96816

Dr. Neil McAlister
Office of Chief of Naval Operations
Command and Control
OP-094H
Washington, D. C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Dr. A. F. Norcio
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, DC 20375-5000

Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D. C. 20350

Mr. John Davis
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Department
Code N-71
Naval Training Equipment Center
Orlando, FL 32813

Mr. Norm Beck
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Engineering
Code 441
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. H. Talkington
Engineering & Computer Science
Code 09
Naval Ocean Systems Center
San Diego, CA 92152

Department of the Navy

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Ross Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Dr. L. Chmura
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Letsky
Office of the Chief of Naval Operations (OP-01B7)
Washington, D.C. 20350

Professor Douglas E. Hunter
Defense Intelligence College
Washington, D.C. 20374

CDR C. Hutchins
Code 55
Naval Postgraduate School
Monterey, CA 93940

Human Factors Technology Administration
Office of Naval Technology
Code MAT 0722
800 North Quincy Street
Arlington, VA 22217-5000

CDR Tom Jones
Naval Air Systems Command
Human Factors Programs
NAVAIR 330J
Washington, D.C. 20361

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D. C. 20361

Mr. Philip Andrews
Naval Sea Systems Command
NAVSEA 61R
Washington, D. C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 81323
Washington, D. C. 20360

Aircrew Systems Branch
Systems Engineering Test Directorate
U.S. Naval Test Center
Patuxent River, MD 20670

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

CAPT Robert Biersner
Naval Biodynamics Laboratory
Michoud Station
Box 29407
New Orleans, LA 70189

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD

Dr. George Moeller
Human Factors Engineering Branch
Naval Submarine Base
Submarine Medical Research Lab.
Groton, CT 06340

Department of the Navy

Head
Aerospace Psychology Department
Naval Aerospace Medical Research Lab
Pensacola, FL 32508

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dr. Jerry Tobias
Auditory Research Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Dr. Robert Blanchard
Code 17
Navy Personnel Research and Development Center
San Diego, CA 92152-6800

LCDR T. Singer
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Stephen Merriman
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

LT. Dennis McBride
Human Factors Branch
Pacific Missle Test Center
Point Mugu, CA 93042

Dr. Kenneth L. Davis
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

LCDR R. Carter
Office of Chief on Naval Operations (OP-01B)
Washington, D.C. 20350

Dean of the Academic Departments
U. S. Naval Academy
Annapolis, MD 21402

CDR W. Moroney
Naval Air Development Center
Code 602
Warminster, PA 18974

Human Factor Engineering Branch
Naval Ship Research and Development Center, Annapolis Division
Annapolis, MD 21402

Dr. Harry Crisp
Code N 51
Combat Systems Department
Naval Surface Weapons Center
Dahlgren, VA 22448

Mr. John Quirk
Naval Coastal Systems Laboratory
Code 712
Panama City, FL 32401

Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Dr. Charles Holland
Office of Naval Research Branch Office
London
Box 39
EPO New York 09510

Dr. Rabinder N. Madan
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Eugene E. Gloye
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

Dr. David Mizell
ONR Detachment
1030 Green Street
Pasadena, CA 91106-2485

Dr. Glen Allgaier
Artificial Intelligence Branch
Code 444
Naval Electronics Ocean System Center
San Diego, CA 92152

Dr. Steve Sacks
Naval Electronics Systems Command
Code 61R
Washington, D.C. 20363-5100

Dr. Sherman Gee
Command and Control Technology,(MAT 0721)
Office of Naval Technology,
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Robert A. Fleming
Human Factors Support Group
Naval Personnel Research & Development Ctr.
1411 South Fern Street
Arlington, VA 22202

Dr. Dick Kelly
Human Factors Division, Code 17
Naval Personnel Research & Development
Center
San Diego, CA 92152-6800

Department of the Army

Dr. Edgar M. Johnson
Technical Director
U.S. Army Research Institute
Alexandria, VA 22333-5600

Technical Director
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005

Director, Organizations and Systems
Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Robert M. Sasmor
Director, Basic Research
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Department of the Air Force

Dr. Kenneth R. Boff
AF AMRL/HE
Wright-Patterson AFB, OH 45433

Dr. A. Fregly
U. S. Air Force Office of
Scientific Research
Life Science Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332-6448

Mr. Charles Bates, Director
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Dr. R. K. Dismukes
Associate Director for Life
Sciences
AFSOR
Bolling AFB
Washington, D.C. 20032-6448

Foreign Addresses

Dr. A. D. Baddeley
Director, Applied Psychology
Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF England

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Tech. Estab.
Teddington, Middlesex
TW11 OLN
England

Other Government Agencies

Dr. M. C. Montemerlo
Information Sciences &
Human Factors Code RC
NASA HQS
Washington, D.C. 20546

Dr. Alan Leshner
Deputy Division Director
Division of Behavioral and
Neural Sciences
National Science Foundation
1800 G. Street, N.W.
Washington, D.C. 20550

Defense Technical Information
Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

Dr. Clinton Kelly
Defense Advanced Research
Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Other Organizations

Dr. Harry Snyder
Dept. of Industrial Engineering
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061

Dr. Amos Tversky
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Dr. J. O. Chinnis, Jr.
Decision Science Consortium,
Inc.
7700 Leesburg Pike
Suite 421
Falls Church, VA 22043

Dr. T. B. Sheridan
Dept. of Mechanical
Engineering
Massachusetts Institute of
Technology
Cambridge, MA 02139

Dr. Daniel Kahneman
The University of British
Department of Psychology
#154-2053 Main Mall
Vancouver, British Columbia
Canada V6T 1Y7

Dr. Stanley Deutsch
NAS-National Research Council
(COHF)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Dr. Meredith P. Crawford
American Psychological
Association
Office of Educational Affairs
1200 17th Street N.W.
Washington, D.C. 20036

Dr. Deborah Boehm-Davis
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. Paul E. Lehner
PAR Technology Corp.
7926 Jones Branch Drive
Suite 170
McLean, VA 22102

Other Organizations

Dr. Babur M. Pulat
Department of Industrial Engineering
North Carolina A&T State University
Greensboro, NC 27411

Dr. Stanley N. Roscoe
University of Colorado
Boulder, CO 80309

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D. C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Dr. Robert Wherry
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Co.
St. Louis Division
Box 516
St. Louis, MO 63166

Dr. Lola L. Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Mr. Joseph G. Whol
Alphatech, Inc.
3 New England Executive Park
Burlington, MA 10803

Dr. Marvin Cohen
Decision Science Consortium, Inc.
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. William R. Utal
NOSC-Hawaii
Box 997
Kailua, HI 96734

Dr. William B. Rouse
School of Industrial and Systems
Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Ms. Denise Benel
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Andrew P. Sage
Assoc. V. P. for Academic Affairs
George Mason University
4400 University Drive
Fairfax, VA 22030

Other Organizations

Dr. Richard Pew
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Hillel Einhorn
Graduate School of Business
University of Chicago
1101 E. 58th Street
Chicago, IL 60637

Dr. Douglas Towne
University of Southern California
Behavioral Technology Lab
1845 South Elena Avenue, Fourth Floor
Redondo Beach, CA 90277

Dr. James P. Jenkins
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. John Payne
Graduate School of Business
  Administration
Duke University
Durham, NC 27706

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Azad Madni
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Tomaso Poggio
Massachusetts Institute of Tech.
Center for Biological Information
 Processing
Cambridge, MA 02139

Dr. Robert A. Hummel
New York University
Courant Inst. of Mathematical
Sciences
251 Mercer Street
New York, New York 10012

Dr. H. McI. Parsons
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Paul Solvic
Decision Research
1201 Oak Street
Eugene, OR 97401

Other Organizations

Dr. David Castanon
ALPHATECH, Inc.
111 Middlesex Turnpike
Burlington, MA 01803

Dr. E. Douglas Jensen
Carnegie-Mellon University
Computer Science Dept.
Pittsburgh, PA 15213

Dr. David Noble
Engineering Research Assoc.
8616 Westwood Center Drive
McLean, VA 22180

Dr. Bruce Hamill
The Johns Hopkins Univ.
Applied Physics Lab
Laurel, MD 20707

Dr. A. Ephremides
University of Maryland
Electrical Engineering Dept.
College Park, MD 20742

Dr. Michael Athans
Massachusetts Inst. of Technology
Lab Information & Decision Systems
Cambridge, MA 01803

Dr. Leonard Adelman
PAR Technology Corp.
7962 Jones Branch Drive
Suite 170
McLean, VA 22102

Dr. Baruch Fischhoff
Perceptronics, Inc.
6271 Variel Ave.
Woodland Hills, CA 91367

END

FILMED

11-85

DTIC